Diversity of MIMO Multihop Relay Channels—Part I: Amplify-and-Forward

Sheng Yang and Jean-Claude Belfiore

Abstract

In this two-part paper, we consider the multiantenna multihop relay channels in which the source signal arrives at the destination through N independent relaying hops in series. The main concern of this work is to design relaying strategies that utilize efficiently the relays in such a way that the diversity is maximized. In part I, we focus on the amplify-and-forward (AF) strategy with which the relays simply scale the received signal and retransmit it. More specifically, we characterize the diversity-multiplexing tradeoff (DMT) of the AF scheme in a general multihop channel with arbitrary number of antennas and arbitrary number of hops. The DMT is in closed-form expression as a function of the number of antennas at each node. First, we provide some basic results on the DMT of the general Rayleigh product channels. It turns out that these results have very simple and intuitive interpretation. Then, the results are applied to the AF multihop channels which is shown to be equivalent to the Rayleigh product channel, in the DMT sense. Finally, the project-and-forward (PF) scheme, a variant of the AF scheme, is proposed. We show that the PF scheme has the same DMT as the AF scheme, while the PF can have significant power gain over the AF scheme in some cases. In part II, we will derive the upper bound on the diversity of the multihop channels and show that it can be achieved by partitioning the multihop channel into AF subchannels.

Index Terms

Multihop, multiple-input multiple output (MIMO), relay channel, amplify-and-forward (AF), diversity-multiplexing tradeoff (DMT).

Diversity of MIMO Multihop Relay Channels—Part I: Amplify-and-Forward

I. INTRODUCTION AND PROBLEM DESCRIPTION

Wireless relaying systems have lots of advantages over traditional direct transmission systems. For example, the periphery can be extended by the relays and the coverage of the existing network can be improved. Using relays can also shorten the point to point transmission distance, which results in lower power (interference) level or in higher throughput. Furthermore, all these benefits can be realized in a more flexible, easier and cheaper to deploy network.

Recently, there has been a boosting interest in the cooperative diversity with which the spatial diversity is exploited through distributed relays. Since the work of Sendonaris *et al.* [1], [2] that introduced the notion of cooperative diversity, a number of relaying protocols have been proposed (see, *e.g.*, [3]–[10]). Most of the previous works consider the single-antenna two-hop relay channel where the source signal is able to arrive at the destination through at most two hops, *i.e.*, the source-relay hop and relay-destination hop. In an N-relay channel, it is shown that a diversity order of $N+1$ (respectively, N) can be achieved with (respectively, without) the direct source-destination link.

In this work, we consider the MIMO multihop channel model without direct source-destination link. That is, the source signal arrives at the destination through N independent relaying hops in series. In the two-hop case, our model is reduced to the model studied by Jing and Hassibi [6]. The central concern of our work is to design relaying strategies that utilize efficiently the relays in such a way that the diversity is maximized. In part I, we focus on the amplify-and-forward (AF) strategy with which the relays simply scale the received signal and retransmit it. The main contributions of this paper are as follows.

1) First, we obtain the diversity-multiplexing tradeoff (DMT) of the Rayleigh product channel, whose channel matrix is a product of independent Gaussian matrices. It turns out that each Rayleigh product channel belongs to an equivalent class that is uniquely represented by the so-called *minimal form*. Furthermore, based on the closed-form expression of the DMT, we derive a recursive DMT characterization that have very simple and intuitive interpretation.

2) Then, it is shown that the AF multihop channel is actually equivalent to the Rayleigh product channel. We can thus identify the two channels and all previously established results apply to the multihop channel. Therefore, the diversity properties of the AF multihop channel in terms of the number of hops and the number of antennas in each node are completely characterized. We also propose the project-and-forward (PF) scheme, a variant of the AF scheme, in the case where full antenna cooperation is possible. It is shown that, although the PF scheme has the same DMT as the AF scheme, the PF can have significant power gain over the AF scheme in some cases.

3) Finally, it is pointed out that using less relaying antennas improve the power gain by avoiding the *hardening* of relayed noise, a particular phenomenon in the AF multihop channel. And reducing the number of transmit antennas can lower significantly the coding delay and decoding complexity. The vertical channel reduction result gives exactly the minimum number of antennas we need at each node to keep the same DMT.

In part II of this paper, we will derive an upper bound on the diversity of the multihop channels and show that the AF scheme is not optimal in general. Then, we will proposed both distributed and non-distributed schemes that achieve the upper bound. The main idea is to partition the multihop channel into AF subchannels.

The rest of part I is organized as follows. Section II presents the channel model and the AF scheme with some basic assumptions. The Rayleigh product channel is introduced and studied in section III. Results concerning the AF and PF schemes are collected in section IV. In section V, numerical results on some typical scenarios are shown. Finally, we draw a brief conclusion in section VI. For fluidity of the presentation, all demonstrations of proofs are delayed to the appendices.

In this paper, we use boldface lower case letters $\boldsymbol{v}$ to denote vectors, boldface capital letters $\boldsymbol{M}$ to denote matrices. $\mathcal{CN}$ represents the complex Gaussian random variable. $[\cdot]^{\mathsf{T}}, [\cdot]^{\dagger}$ respectively denote the matrix transposition and conjugated transposition operations. $\|\cdot\|$ is the vector norm. $(x)^{+}$ means $\max(0, x)$. $\mathrm{Det}(\boldsymbol{M})$ is the absolute value of the determinant $\det(\boldsymbol{M})$. The square root $\sqrt{\boldsymbol{P}}$ of a positive semi-definite matrix $\boldsymbol{P}$ is defined as a positive semi-definite matrix such that $\boldsymbol{P} = \sqrt{\boldsymbol{P}}(\sqrt{\boldsymbol{P}})^{\dagger}$. The ordered eigenvalues of a positive semi-definite matrix $\boldsymbol{P}$ are denoted
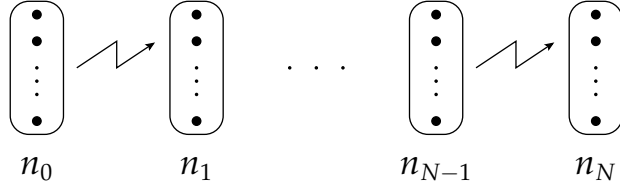
Fig. 1. A MIMO multihop relay channel.

by $\boldsymbol{\lambda}(\boldsymbol{P})$ or $\boldsymbol{\mu}(\boldsymbol{P})$. We define $\boldsymbol{\alpha}(\boldsymbol{P})$ and $\boldsymbol{\beta}(\boldsymbol{P})$ by

$$\alpha_i(\boldsymbol{P}) \triangleq -\log \lambda_i(\boldsymbol{P}) / \log \mathsf{SNR} \quad \text{and} \quad \beta_i(\boldsymbol{P}) \triangleq -\log \mu_i(\boldsymbol{P}) / \log \mathsf{SNR}.$$

And we call them the *eigen-exponents* of $\boldsymbol{P}$, with a slight abuse of terminology. We drop the arguments of $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ when confusion is not likely. For any quantity q,

$$q \doteq \mathsf{SNR}^a \quad \text{means} \quad \lim_{\mathsf{SNR} \to \infty} \frac{\log q}{\log \mathsf{SNR}} = a$$

and similarly for $\dot{\leq}$ and $\dot{\geq}$. The tilde notation $\tilde{\boldsymbol{n}}$ is used to denote the (increasing) ordered version of $\boldsymbol{n}$. Let $\boldsymbol{m}$ and $\boldsymbol{n}$ be two vectors of same length L, then $\boldsymbol{m} \preceq \boldsymbol{n}$ means $\tilde{m}_i \leq \tilde{n}_i, \forall i$.

II. SYSTEM MODEL

A. Channel Model

The considered N-hop relay channel model is illustrated in Fig. 1, where there are one source (node #0), one destination (node #N), and $N-1$ clusters of intermediate relays. Each cluster is logically seen as a node (node #1 to node #$N-1$) that is equipped with multiple antennas (n_i antennas for node #i). We assume that node #i can only hear node #$i-1$. Mathematically, we have

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{x}_{i-1} + \boldsymbol{z}_i$$

where $\boldsymbol{H}_i \in \mathbb{C}^{n_i \times n_{i-1}}$ is the channel between node #$i-1$ and node #i; $\boldsymbol{x}_i, \boldsymbol{y}_i \in \mathbb{C}^{n_i \times 1}$ is the transmitted and received signal at node #i; $\boldsymbol{z} \in \mathbb{C}^{n_i \times 1} \in \mathcal{CN}(0, \mathbf{I})$ is the additive white Gaussian noise at node #i. The channels $\boldsymbol{H}_i$'s are independent and modeled as Rayleigh quasi-static channels, *i.e.*, the entries of $\boldsymbol{H}_i$ are i.i.d. $\mathcal{CN}(0, 1)$ distributed and do not change during the
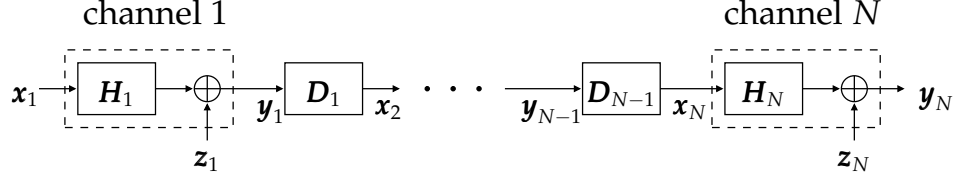
Fig. 2. Amplify-and-forward strategy for multihop channels.

transmission of a data frame. For simplicity, it is assumed that the intermediate nodes work in full-duplex[1] mode and all transmitting nodes are subject to the same short-term power constraint

$$\mathbb{E}\{\|\boldsymbol{x}_i\|^2\} \leq \mathsf{SNR}, \quad \forall\, i \tag{1}$$

where the expectation is taken on the noises. All terminals are supposed to have full channel state information (CSI) at the receiver and no CSI at the transmitter. From now on, we denote the channel as a $(n_0, n_1, \ldots, n_N)$ multihop channel.

B. Amplify-and-Forward Protocol

The AF strategy is described as follows. At each node, the received signal of each antenna is normalized to the same power level and then retransmitted. As shown in Fig. 2, the signal model is

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{x}_i + \boldsymbol{z}_i,$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{D}_i \boldsymbol{y}_i$$

where the transmitted signal $\boldsymbol{x}_i$ has the short-term power constraint

$$\mathbb{E}\left(|\boldsymbol{x}_i[j]|^2\right) \leq \frac{\mathsf{SNR}}{n_i};$$

the scaling matrix $\boldsymbol{D}_i \in \mathbb{C}^{n_i \times n_i}$ is diagonal with the normalization factors[2]

$$\boldsymbol{D}_i[j,j] = \sqrt{\frac{1}{\frac{\mathsf{SNR}}{n_{i-1}}\left(\sum_{k=1}^{n_{i-1}} |\boldsymbol{H}_i[j,k]|^2\right) + 1}} \cdot \sqrt{\frac{\mathsf{SNR}}{n_i}}. \tag{2}$$

---

[1]The assumption is merely for simplicity of notation. As one can easily verify, since no cross-talk between different channels, the half-duplex constraint is directly translated to a reduction of degrees of freedom by a factor of two and does not impact the relaying strategy. This is achieved by letting all even-numbered (respectively, odd-numbered) nodes transmit (respective, receive) in even-numbered time slot and received (respective, transmit) in odd-numbered time slots.

[2]In the case where long-term power constraint is imposed, we simply replace the channel coefficients $|\boldsymbol{H}_i[j,k]|$ in (2) by 1's.

C. Diversity-Multiplexing Tradeoff

In this paper, we use the diversity-multiplexing tradeoff (DMT) as the performance measure.

Definition 1 (Multiplexing and diversity gains [11]): The *multiplexing gain* r and *diversity gain* d of a fading channel are defined by

$$r \triangleq \lim_{\mathsf{SNR} \to \infty} \frac{R(\mathsf{SNR})}{\log \mathsf{SNR}} \quad \text{and} \quad d \triangleq - \lim_{\mathsf{SNR} \to \infty} \frac{\log P_{\text{out}}(\mathsf{SNR}, R)}{\log \mathsf{SNR}}$$

where $R(\mathsf{SNR})$ is the target data rate and $P_{\text{out}}(\mathsf{SNR}, R)$ is the outage probability for a target rate R. A more compact form is

$$P_{\text{out}}(\mathsf{SNR}, r \log \mathsf{SNR}) \doteq \mathsf{SNR}^{-d}. \tag{3}$$

Note that in the definition we use the outage probability instead of the error probability, since it is shown in [11] that the error probability is dominated by the outage probability in the high SNR regime and that the thus defined DMT is the best that we can achieve with any coding scheme.

Lemma 1: The DMT of a $n_{\text{t}} \times n_{\text{r}}$ Rayleigh channel is a piecewise-linear function connecting the points $(k, d(k))$, $k = 0, 1, \ldots, \min(n_{\text{t}}, n_{\text{r}})$, where

$$d(k) = (n_{\text{t}} - k)(n_{\text{r}} - k).$$

III. THE RAYLEIGH PRODUCT CHANNEL

As it is shown in the next section, the AF multihop channels are intimately related to a more general Rayleigh product channel defined below. In this section, we investigate the Rayleigh product channel and provides some basic results on the diversity. Let us begin by the following definitions.

Definition 2 (Rayleigh product channel): Let $\boldsymbol{H}_i \in \mathbb{C}^{n_{i-1} \times n_i}$, $i = 1, 2, \ldots, N$, be N independent complex Gaussian matrices with i.i.d. zero mean unit variance entries. A $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel is a $n_N \times n_0$ MIMO channel defined by

$$\boldsymbol{y} = \sqrt{\frac{\mathsf{SNR}}{n_1 \cdots n_N}} \boldsymbol{\Pi} \boldsymbol{x} + \boldsymbol{z} \tag{4}$$

where $\boldsymbol{\Pi} \triangleq \boldsymbol{H}_1 \boldsymbol{H}_2 \cdots \boldsymbol{H}_N$; $\boldsymbol{x}$ is the transmitted signal with power constraint $\mathbb{E}(\|\boldsymbol{x}\|^2) \leq n_N$; $\boldsymbol{z} \in \mathbb{C}^{n_0 \times 1} \sim \mathcal{CN}(0, \boldsymbol{I})$ is the additive white Gaussian noise; SNR is the receive signal-to-noise ratio (SNR) per receive antenna.

Definition 3 (Exponential equivalence): Two channels are said to be *exponentially equivalent* or *equivalent* if their eigen-exponents have the same asymptotical joint pdf.

Let $\tilde{\boldsymbol{n}}$ be the ordered version of $\boldsymbol{n}$ with $\tilde{n}_N \geq \tilde{n}_{N-1} \geq \cdots \geq \tilde{n}_0$.

Definition 4 (Reduction of Rayleigh product channel): A $(m_0, m_1, \ldots, m_k)$ Rayleigh product channel is said to be a *reduction* of a $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel if 1) they are equivalent, 2) $k \leq N$, and 3) $(m_0, m_1, \ldots, m_k) \preceq (\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_k)$. In particular, if $k = N$, then it is called a *vertical reduction*. Similarly, if $\tilde{m}_i = \tilde{n}_i, \ \forall\, i \in [0, k]$, it is a *horizontal reduction*.

Definition 5 (Minimal form): $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_{N^*})$ is said to be a *minimal form* if no reduction other than itself exists. Similarly, it is called a *minimal vertical form* (respectively, *minimal horizontal form*) if no vertical (respectively, horizontal) reduction other than itself exists. A channel is said to have *order* N^* if its minimal form is of length $N^* + 1$.

A. Joint PDF of the Eigen-exponents of $\boldsymbol{\Pi\Pi}^\dagger$

Theorem 1: Let us denote the non-zero ordered eigenvalues of $\boldsymbol{\Pi\Pi}^\dagger$ by $\lambda_1 \geq \cdots \geq \lambda_{n_{\min}} > 0$ with $n_{\min} \triangleq \min\limits_{i=0,\ldots,N} n_i$. Then, the joint pdf of the eigen-exponents $\boldsymbol{\alpha}$ satisfies

$$p(\boldsymbol{\alpha}) \doteq \begin{cases} \mathsf{SNR}^{-E(\boldsymbol{\alpha})}, & \text{for } 0 \leq \alpha_1 \leq \ldots \leq \alpha_{n_{\min}}, \\ \mathsf{SNR}^{-\infty}, & \text{otherwise} \end{cases} \tag{5}$$

where

$$E(\boldsymbol{\alpha}) \triangleq \sum_{i=1}^{n_{\min}} c_i \alpha_i \tag{6}$$

with

$$c_i \triangleq 1 - i + \min_{k=1,\ldots,N} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - i}{k} \right\rfloor, \quad i = 1, \ldots, n_{\min}. \tag{7}$$

By definition, $n_{\min} = \tilde{n}_0$ and we interchange the notations depending on the context. From the theorem, we can see that the asymptotical eigen-exponents distribution depends only on $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_N)$, the ordered version of $(n_0, n_1, \ldots, n_N)$. For example, a $(3, 1, 4, 2)$ channel is equivalent to a $(1, 2, 3, 4)$ channel, in the eigen-exponent sense.

Theorem 2: A $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel can be reduced to a $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_k)$ channel if and only if

$$k(\tilde{n}_{k+1} + 1) \geq \sum_{l=0}^{k} \tilde{n}_l. \tag{8}$$

In particular, it can be reduced to a Rayleigh channel if and only if

$$\tilde{n}_2 + 1 \geq \tilde{n}_0 + \tilde{n}_1. \tag{9}$$

This theorem implies that $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_{N^*})$ is a minimal form if there exists no $k < N^*$ such that (8) is satisfied. One can also verify that if $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_{N^*})$ is a minimal horizontal form of $(n_0, n_1, \ldots, n_N)$, then 1) it is also a minimal form; and 2) the minimal vertical form is $(\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_{N^*}, \bar{n}, \ldots, \bar{n})$ where

$$\bar{n} = \left\lceil \frac{\sum_{l=0}^{N^*} \tilde{n}_i}{N^*} - 1 \right\rceil. \tag{10}$$

Furthermore, note that the order N^* is upper-bounded by $\tilde{n}_0$ because (8) is always satisfied with $k = \tilde{n}_0$. In other words, the length of the minimal form is bounded by $\tilde{n}_0 + 1$. In particular, the minimal form of a $(1, n_1, \ldots, n_N)$ Rayleigh product channel is always $(1, n_1)$, *i.e.*, a $1 \times \tilde{n}_1$ or $\tilde{n}_1 \times 1$ Rayleigh channel.

Theorem 3: Two Rayleigh product channels are equivalent if and only if they have the same minimal form.

From this theorem, we deduce that the class of exponential equivalence is *uniquely* identified by the minimal form. Therefore, N^* can also be defined as the order of the class.

B. Characterization of the Diversity-Multiplexing Tradeoff

From theorem 1, we can derive the DMT of a Rayleigh product channel.

Theorem 4 (Direct characterization): The DMT of a Rayleigh product channel $(n_0, n_1, \ldots, n_N)$ is a piecewise-linear function connecting the points $(k, d(k))$, $k = 0, 1, \ldots, n_{\min}$, where

$$d(k) = \sum_{i=k+1}^{n_{\min}} c_i \tag{11}$$

with c_i defined by (7).

Since the DMT is a bijection of the coefficients c_i's, all results obtained previously apply to the DMT and two Rayleigh product channels are equivalent if and only if they have the same DMT. Hence, the exponential equivalence class is also the DMT-equivalence class. However, unlike the eigen-exponent, the DMT provides an insight on the diversity performance of a channel (or a scheme) for different multiplexing gain. Note that, despite the closed-form nature of the characterization (11), it is lack of intuition. That is why we search for an alternative characterization.
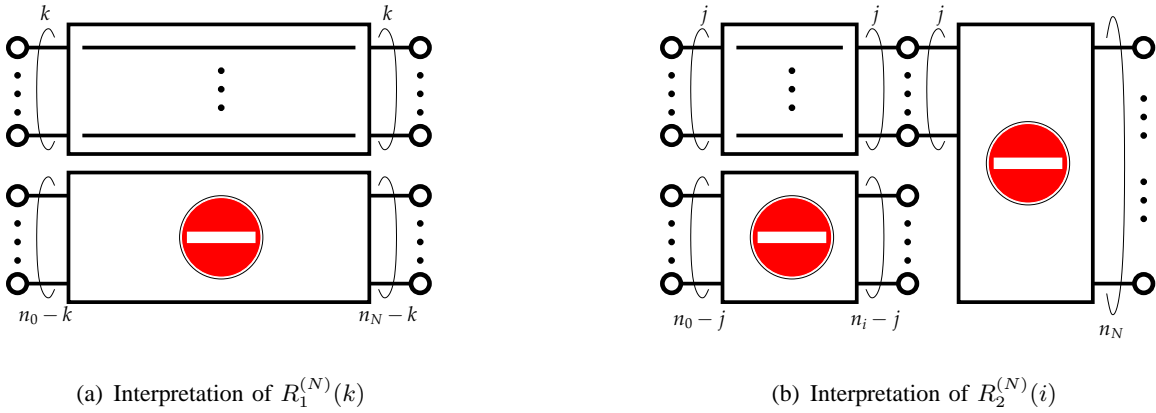
(a) Interpretation of $R_1^{(N)}(k)$

(b) Interpretation of $R_2^{(N)}(i)$

Fig. 3. Interpretations of the DMT of the Rayleigh product channel.

Theorem 5 (Recursive characterization): The DMT $d(k)$ defined in (11) can be alternatively characterized by

$$R_1^{(N)}(k) \quad : \quad d_{(n_0,\dots,n_N)}(k) = d_{(n_0-k,\dots,n_N-k)}(0), \quad \forall k; \tag{12}$$

$$R_2^{(N)}(i) \quad : \quad d_{(n_0,\dots,n_N)}(0) = \min_{j \geq 0} d_{(n_0,\dots,n_i)}(j) + d_{(j,n_{i+1},\dots,n_N)}(0), \quad \forall i; \tag{13}$$

$$R_3^{(N)}(i,k) \quad : \quad d_{(n_0,\dots,n_N)}(k) = \min_{j \geq k} d_{(n_0,\dots,n_i)}(j) + d_{(j,n_{i+1},\dots,n_N)}(k), \quad \forall i,k. \tag{14}$$

The recursive characterization has an intuitive interpretation as follows. Let us consider k as a "*network flow*" between the source and the destination and $d(k)$ as the *minimum "cost"* to limit the flow to k (the flow-k event). In particular, the maximum diversity $d(0)$ can be seen as the "*disconnection cost*". First, $R_1(k)$ says that the most efficient way to limit the flow to k is to keep a $(k, k, \dots, k)$ channel fully connected and to disconnect the $(n_0 - k, n_1 - k, \dots, n_N - k)$ residual channel, as shown in Fig. 3(a). Then, $R_2(i)$ suggests that in order to disconnect a $(n_0, n_1, \dots, n_N)$ channel, if we allow for j flows from the source to some node i, then the $(j, n_{i+1}, \dots, n_N)$ channel from the j "ends" of the flows at node i to the destination must be disconnected. The idea is shown in Fig. 3(b). Obviously, the most efficient way is such that the total cost is minimized with respect to j. This interpretation sheds lights on the typical outage event of the Rayleigh product channel. In the trivial case of $N = 1$ (the Rayleigh channel), there is only one subchannel. The typical and only way for the channel to be in outage is that all the paths are bad, *i.e.*, the disconnection cost is $\tilde{n}_0 \times \tilde{n}_1$. In the non-trivial cases, there are more than

one subchannels and thus the typical outage event is not necessarily for one of the subchannels being totally bad. The *mismatch* of two partially bad subchannels can also cause outage. In a more general way, the flow-k event takes place when both the flow-j event in the $(n_0, \ldots, n_i)$ channel and the flow-k event in the $(j, n_{i+1}, \ldots, n_k)$ channel happen at the same time. We can verify that $(R_1(k), R_3(i, k))$ is equivalent to $(R_1(k), R_2(i))$. Note that the DMT is completely characterized by these relations in a recursive manner.

The following corollaries conclude some properties of the DMT of the Rayleigh product channel.

Corollary 1 (Monotonicity): The DMT is monotonic in the following senses :

1) if $(n_{1,0}, n_{1,1}, \ldots, n_{1,N}) \succeq (n_{2,0}, n_{2,1}, \ldots, n_{2,N})$, then

$$d_{(n_{1,0}, \ldots, n_{1,N})}(r) \geq d_{(n_{2,0}, \ldots, n_{2,N})}(r), \quad \forall\, r;$$

2) if $\{n_{1,0}, n_{1,1}, \ldots, n_{1,N_1}\} \supseteq \{n_{2,0}, n_{2,1}, \ldots, n_{2,N_2}\}$, then

$$d_{(n_{1,0}, \ldots, n_{1,N_1})}(r) \leq d_{(n_{2,0}, \ldots, n_{2,N_2})}(r), \quad \forall\, r.$$

Corollary 2: Let us define

$$
p_k \triangleq
\begin{cases}
\tilde{n}_0 & k = 0, \\
\sum_{l=0}^{k} \tilde{n}_l - k\tilde{n}_{k+1} & k = 1, \ldots, N-1, \\
-\infty & k = N.
\end{cases}
\tag{15}
$$

Then,

$$d_{(n_0, \ldots, n_N)}(r) = d_{(\tilde{n}_0, \ldots, \tilde{n}_k)}(r), \quad \text{for } r \geq p_k.$$

While corollary 1 implies that $d(r) \leq d_{(\tilde{n}_0, \ldots, \tilde{n}_k)}(r)$ in a general way, corollary 2 states precisely that $d(r)$ coincides with $d_{(\tilde{n}_0, \ldots, \tilde{n}_k)}(r)$ for $r \geq p_k$.

Corollary 3 (Upper bound and lower bound):

$$\frac{\tilde{n}_0 \tilde{n}_1}{2} < d(0) \leq \tilde{n}_0 \tilde{n}_1$$

where $d(0)$ is known as the maximum diversity gain.

From (7) and (11), the upper bound is obtained by setting $\tilde{n}_2$ large enough and the lower bound is obtain by setting $\tilde{n}_2 = \ldots = \tilde{n}_N$. This corollary implies that the diversity of a Rayleigh product channel can always be written as $d(0) = a\tilde{n}_0 \tilde{n}_1$ with $a \in (0.5, 1]$. Hence, the diversity

"bottleneck" of the Rayleigh product channel $\boldsymbol{\Pi}$ is not necessarily one of the subchannels $\boldsymbol{H}_i$, but rather the virtual $\tilde{n}_0 \times \tilde{n}_1$ Rayleigh channel. On the other hand, the maximum diversity gain is always strictly larger than $\frac{\tilde{n}_0 \tilde{n}_1}{2}$, independent of the value N. In order to illuminate the impact of N on the DMT, let us consider the symmetric case.

Corollary 4 (Symmetric Rayleigh product channels): When $n_0 = n_1 = \ldots = n_N = n$, we have

$$d(k) = \frac{(n-k)(n+1-k)}{2} + \frac{a(k)}{2}((a(k)-1)N + 2b(k)) \tag{16}$$

where $a(k) \triangleq \left\lfloor \frac{n-k}{N} \right\rfloor$ and $b(k) \triangleq (n-k) \bmod N$.

In the symmetric case, on one hand, we observe that the DMT degrades with N. On the other hand, from (16), the degradation stops at $N = n$ and we have

$$d(k) = \frac{(n-k)(n+1-k)}{2}$$

for $N \geq n$. This can also be deduced from theorem 2 applying which we get that the order of all symmetric Rayleigh product channel with $N > n$ is $N^* = n$. Therefore, we lose less than half of the diversity gain due to the product of Rayleigh MIMO channels, in contrast to the intuition that the maximum diversity gain could degrade to 1 with $N \to \infty$. As an example, in Fig. 4, we show the DMT of the 2×2 and 5×5 Rayleigh product channels with different values of N.

C. General Rayleigh Product Channel

In fact, we can define a more general Rayleigh product channel as

$$\boldsymbol{\Pi}_g \triangleq \boldsymbol{H}_1 \boldsymbol{T}_{1,2} \boldsymbol{H}_2 \cdots \boldsymbol{H}_{N-1} \boldsymbol{T}_{N-1,N} \boldsymbol{H}_N. \tag{17}$$

Theorem 6: The general Rayleigh product channel is equivalent to

1) a $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel, if all the matrices $\boldsymbol{T}_{i,i+1}$'s are square and their singular values satisfy $\sigma_j(\boldsymbol{T}_{i,i+1}) \doteq \mathsf{SNR}^0$, $\forall i, j$;

2) a $(n_0, n'_1, \ldots, n'_{N-1}, n_N)$ Rayleigh product channel, with n'_i being the rank of the matrix $\boldsymbol{T}_{i,i+1}$, if the matrices $\boldsymbol{T}_{i,i+1}$'s are constant.

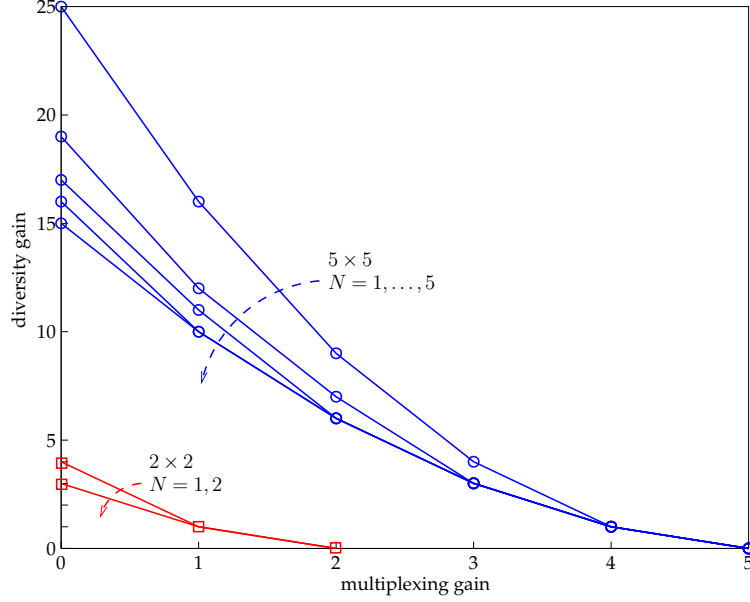Therefore, the results obtained previously for the Rayleigh product channel can be applied to the general one.

Fig. 4. Diversity-multiplexing tradeoff of 2×2 and 5×5 symmetric Rayleigh product channels.

IV. AMPLIFY-AND-FORWARD MULTIHOP CHANNELS

Using the results from the previous section, we are going to analyze the performance of the AF scheme presented in section II, in terms of the DMT.

A. Equivalence to the Rayleigh Product Channel

With the AF scheme, the end-to-end equivalent MIMO channel is

$$\boldsymbol{y}_N = \left(\prod_{i=1}^{N} \boldsymbol{D}_i \boldsymbol{H}_i \right) \boldsymbol{x}_1 + \sum_{j=1}^{N} \left(\prod_{i=j}^{N} \boldsymbol{H}_{i+1} \boldsymbol{D}_i \right) \boldsymbol{z}_j \tag{18}$$

where for the sake of simplicity, we define $\prod_{i=1}^{N} \boldsymbol{A}_i \triangleq \boldsymbol{A}_N \cdots \boldsymbol{A}_1$ for any matrices $\boldsymbol{A}_i$'s; $\boldsymbol{H}_{N+1} \triangleq \mathbf{I}$ and $\boldsymbol{D}_N \triangleq \mathbf{I}$. The standard whitened form of this channel is

$$\boldsymbol{y} = \sqrt{\boldsymbol{R}} \left(\prod_{i=1}^{N} \boldsymbol{D}_i \boldsymbol{H}_i \right) \boldsymbol{x}_1 + \boldsymbol{z}$$

where $\boldsymbol{z} \sim \mathcal{CN}(0, \mathbf{I})$ is the whitened version of the noise and $\sqrt{\boldsymbol{R}}$ is the whitening matrix with $\boldsymbol{R}$ the covariance matrix of the noise in (18). Since it can be shown that $\lambda_{\max}(\boldsymbol{R}) \doteq \lambda_{\min}(\boldsymbol{R}) \doteq \mathrm{SNR}^0$, the AF multihop channel is DMT-equivalent to the channel defined by

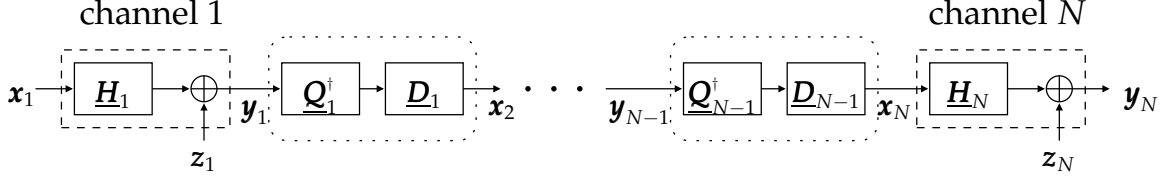$$\boldsymbol{H}_N \boldsymbol{D}_{N-1} \cdots \boldsymbol{H}_2 \boldsymbol{D}_1 \boldsymbol{H}_1,$$

Fig. 5. The project-and-forward scheme.

which is a general Rayleigh product channel defined in (17) if we have $\sigma_j(\boldsymbol{D}_i) \doteq \mathsf{SNR}^0$, $\forall i, j$. To this end, we slightly modify the matrices $\boldsymbol{D}_i$'s and get the new matrices $\hat{\boldsymbol{D}}_i$ with

$$\hat{\boldsymbol{D}}_i[j, j] = \min\left\{\boldsymbol{D}_i[j, j], \kappa\right\}$$

where $0 < \kappa < \infty$ is a constant[3] independent of SNR. Furthermore, it is obvious that the power constraint is still satisfied by replacing $\boldsymbol{D}_i$ with $\hat{\boldsymbol{D}}_i$. Therefore, the multihop channel with the thus defined AF strategy is DMT-equivalent to a $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel, *i.e.*,

$$d^{\mathrm{AF}}(k) = \sum_{i=k+1}^{n_{\min}} c_i.$$

In the rest of the paper, we identify the Rayleigh product channel, the AF multihop channel and the vector $(n_0, n_1, \ldots, n_N)$ when confusion is not likely.

B. A Variant : Project-and-Forward

We propose a new scheme called project-and-forward (PF), as shown in Fig. 5. This scheme can be used only when full antenna cooperation within cluster is possible, that is, all antennas in the same cluster are controlled by a central unit. At the node #i, the received signal is first projected to the signal subspace $\mathcal{S}_i$, spanned by the columns of the channel matrix $\underline{\boldsymbol{H}}_i$. The dimension of $\mathcal{S}_i$ is r_i, the rank of $\underline{\boldsymbol{H}}_i$. After the component-wise normalization, the projected signal is transmitted using r_i (out of n_i) antennas. It is now clear that $\underline{\boldsymbol{H}}_{i+1} \in \mathbb{C}^{n_{i+1} \times r_i}$ is actually composed of the r_i columns of the previously defined $\boldsymbol{H}_{i+1}$, with $r_0 \triangleq n_0$.

More precisely, the $\underline{\boldsymbol{Q}}_i \in \mathbb{C}^{n_i \times r_i}$ is an orthogonal basis of $\mathcal{S}_i$ with $\underline{\boldsymbol{Q}}_i^\dagger \underline{\boldsymbol{Q}}_i = \mathbf{I}$. We can rewrite

$$\underline{\boldsymbol{H}}_i = \underline{\boldsymbol{Q}}_i \underline{\boldsymbol{G}}_i$$

---

[3]The κ is only for theoretical proof and is not used in practice, since we can always set κ a very large constant but independent of SNR. In this case, $\hat{\boldsymbol{D}}_i = \boldsymbol{D}_i$ with probability close to 1 for practical SNR.

with $\underline{\boldsymbol{G}}_i \in \mathbb{C}^{r_i \times r_{i-1}}$. For simplicity, we let $\underline{\boldsymbol{Q}}_i$ be obtained by the QR decomposition of $\underline{\boldsymbol{H}}_i$ if $n_i > r_{i-1}$ and be identity matrix if $n_i \leq r_i$. The main idea of the PF scheme is not to use more antennas than necessary to forward the signal. Since the useful signal lies only in the r_i-dimensional signal subspace, the projection of the received signal provides sufficient statistics and reduces the noise power by a factor $\frac{n_i}{r_i}$. In this case, only r_i antennas are needed to forward the projected signal. Let us define $\boldsymbol{P}_i \triangleq \underline{\boldsymbol{D}}_i \underline{\boldsymbol{Q}}_i^\dagger$. Then, as in the AF case, the PF multihop channel is DMT-equivalent to the channel defined by

$$\boldsymbol{\Pi}_{\mathrm{PF}} = \underline{\boldsymbol{H}}_N \boldsymbol{P}_{N-1} \cdots \underline{\boldsymbol{H}}_2 \boldsymbol{P}_1 \underline{\boldsymbol{H}}_1.$$

The following theorem states that using only r_i out of n_i antennas to forward the projected signal does not incur any loss of diversity, as compared to the AF scheme.

Theorem 7: The PF multihop channel is DMT-equivalent to a $(n_0, n_1, \ldots, n_N)$ Rayleigh product channel.

While the PF and AF have the same diversity gain, the PF outperforms the AF in power gain for two reasons. One reason is, as stated before, that the projection reduces the average noise power. The other reason is that the accumulated noise in the AF case is more substantial than that in the PF case. This is because in the PF case, less relay antennas are used than in the AF case. Since the power of independent noises from different transmit antennas add up at the receiver side, the accumulated noise in the AF case "enjoys" a larger "transmit diversity order" than in the PF case. We call it the *noise hardening* effect. Some examples will be given in the section of numerical results.

C. Practical Issues

1) Space-Time Coding: From the input-output point of view, the multihop channel with AF/PF protocol is merely a linear MIMO fading channel, for which the DMT-achieving space-time codes exist. For example, in [11], a Gaussian code is shown to achieve the DMT of a $n_0 \times n_1$ Rayleigh channel if the code length $l \geq n_0 + n_1 - 1$. This result can easily be extended to a general linear fading channel and one can show that Gaussian coding is DMT-achieving for any fading statistics if l is large enough.

Another family of code construction is based on cyclic division algebra (CDA). These codes have minimum length n_0 and are commonly known as the Perfect codes [12], [13]. They are

DMT-achieving thanks to the so-called non-vanishing determinant (NVD) property. It has been shown that they are approximately universal [13], [14] since they are DMT-achieving for all fading statistics. Therefore, we propose to use the rate-$\tilde{n}_0$ $n_0 \times n_0$ Perfect codes. In this case, the only information that the source need to know is $\tilde{n}_0$.

2) Antenna Reduction: In the AF case, provided the number of total available antennas $(n_0, n_1, \ldots, n_N)$, the vertical reduction result gives an exact number of necessary antennas at each node in the DMT sense. This result can be used to reduce the number of transmit and relay antennas[4]. If Perfect space-time codes are used, reducing the number of transmit antennas n_0 means reducing the coding length, *i.e.*, coding delay and decoding complexity, since the code length is equal to the number of transmit antennas. For instance, only two transmit antennas are needed in a $(4, 2, 2, 2)$ channel. Therefore, instead of using a 4×4 Perfect code the code length of which is 4, one can use the Golden code [15] of length 2 and still achieve the DMT.

In fact, less relay antennas also means less relay signaling (relay probing, synchronization, etc.) overhead especially when different antennas are from different relaying terminals (single-antenna relays). Furthermore, using more relay antennas hardens the relayed noise. This is the same phenomenon as we stated in the PF case. Therefore, the number of relay antennas at each node should be restricted to $\bar{n}$ (defined in (10)), the number given by the vertical reduction.

V. EXAMPLES AND NUMERICAL RESULTS

In this section, we provide some examples of multihop channels and show the performance of AF scheme with simulation results. In all cases, we make the same assumptions as in section II.

A. Horizontal and Vertical Reduction

Outage performances versus the received SNR per node of different multihop channels are shown in Fig. 7. Note that both the $(2, 2)$ and $(2, 2, 2)$ channels are minimal and have diversity order 4 and 3, respectively. The $(3, 2, 2)$ channel can be horizontally reduced to $(2, 2)$ and thus has diversity 4. Similarly, the $(2, 2, 2, 2)$, $(4, 2, 2, 2)$ and $(8, 2, 2, 2)$ channels can be reduced to $(2, 2, 2)$ and have diversity 3. As compared to the $(2, 2, 2, 2)$ channel, the larger number of

---

[4]Reducing the number of receive antennas does not do any good, since more receive antennas always provide larger power gain without increasing the complexity.

transmit antennas in the $(8, 2, 2, 2)$ weakens the fading of the first hop and the performance is close to the $(2, 2, 2)$ channel.

Another example is to illustrate the vertical reduction of multihop channels, as shown in Fig. 8. We first consider the case of a $(1, 4, 1)$ channel. The necessary antenna number $\bar{n}$ is 1 and the minimal vertical form is thus $(1, 1, 1)$. We observe that, although both the $(1, 4, 1)$ and $(1, 1, 1)$ channels have diversity 1, a power gain of 7 dB is obtained at $P_{\text{out}} = 10^{-4}$ by using only one relay antennas out of four, if the AF scheme is used. As stated in section IV-C.2, the gain is due to avoiding the hardening of relayed noise. Then, we consider the $(3, 1, 4, 2)$ channel. The necessary number of antennas $\bar{n}$ is 2 in this case. As shown in Fig. 8, by restricting the number of relay antennas to 2, we have a $(3, 1, 2, 2)$ channel and a gain of 2 dB is observered at $P_{\text{out}} = 10^{-4}$. We can further reduce the number of transmit antennas to 2 to get a $(2, 1, 2, 2)$ channel. Unlike the reduction of relay antennas, the reduction of transmit antennas does not provide any gain because it does not affect the relayed noise. In contrast, it degrades the performance since the first hop $(2, 1)$ is faded more seriously than the original first hop $(3, 1)$. Nevertheless, the $(2, 1, 2, 2)$ channel is still better than the $(3, 1, 4, 2)$ channel and is only 0.7 dB from the $(3, 1, 2, 2)$ channel.

B. Project-and-Forward

In Fig. 9, we compare the PF scheme with the AF scheme for the $(1, 2, 1)$ and $(1, 3, 2)$, respectively. First of all, note that the AF and the PF have the same diversity order, as predicted. Then, a power gain of 8.5 dB (respectively, 6.5 dB) over the AF scheme is obtained by the PF scheme in the $(1, 2, 1)$ (respective, $(1, 3, 2)$ channel). This is due to the maximum ratio combining (MRC) gain in the first hop and to avoiding the relayed noise hardening.

C. Coded Performance

We now study the coded performance of the AF multihop channel. The performance measure is the symbol error rate (SER) versus the received SNR under the maximum likelihood (ML) decoding. We still take the $(3, 1, 4, 2)$ channel as an example. Since $\tilde{n}_0 = 1$, the diagonal algebraic space-time (DAST) code[5] [16] can be used. As shown in Fig. 10, with the DAST code, the symbol error rate performances of in the $(3, 1, 4, 2)$, $(3, 1, 2, 2)$ and $(2, 1, 2, 2)$ channels have exactly the

---

[5]Note that the DAST code is the diagonal version of the rate-one Perfect code proposed in [12].

same behavior as the outage performances of the channels do Fig. 8. Moreover, we can use the Alamouti code [17] for the $(2, 1, 2, 2)$ channel. As we can see in the figure, the Alamouti code outperforms all the DAST codes with minimum delay and minimum decoding complexity. The potential benefits from the vertical reduction are thus highlighted.

D. Multihop vs. Direct Transmission

Finally, we introduce the path loss model [18]

$$\text{SNR}_{\text{received}} \propto \text{distance}^{-\alpha} \text{SNR}_{\text{transmitted}}$$

where α is the path loss factor. We fix the distance from the source to the destination and dispose the relay nodes on the source-destination line with equal distance. Each node contains two antennas. We compare the 2-, 3- and 4-hop channel with the direct transmission (single-hop) channel. the performance measure is the transmitted power gain of the multihop channel over the single-hop channel at certain target outage probability (10^{-3} and 10^{-4}). The path loss factor α takes the typical values [18] 3, 3.5, and 4 for wireless channels. In Fig. 11(a), the total transmission power in the multihop channel is considered. Power gain is obtained for $\alpha = 3.5$ and 4. Then, the transmission power per node is considered in Fig. 11(b). In this case, power gain is obtained for all α and is as high as 11 dB. In practice, the transmission power per node also represents the interference level for other terminals which has a significant impact on the network capacity. In both figures, the power gain is lower at 10^{-4} than at 10^{-3}. This is due to the fact that the direct transmission channel is a 2×2 Rayleigh channel and has diversity 4, while the multihop channel is $(2, 2, \ldots, 2)$ and has diversity 3. And low diversity gain means decreasing power gain with increasing SNR or equivalently, with decreasing outage probability.

VI. CONCLUSION

Perhaps the simplest relaying scheme in the MIMO multihop channel is the Amplify-and-Forward scheme. In part I of this paper, by identifying the AF multihop channel with the so-called Rayleigh product channel, we have obtained the complete characterization of the diversity-multiplexing tradeoff of the AF scheme in a multihop channel with arbitrary number of antennas and hops. The characterization is provided both in direct closed-form and recursive form. Based on the DMT, a number of properties of the AF multihop channel have been derived.

In the second part, we will show that the AF scheme is suboptimal in general, by establishing the diversity upper bound of the multihop channel with any relaying scheme. By partitioning the multihop channel into AF subchannels, we achieve the upper bound with both distributed and non-distributed schemes.

APPENDIX I

PRELIMINARIES

The followings are some preliminary results that are essential to the proofs.

Definition 6 (Wishart Matrix): The $m \times m$ random matrix $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^\dagger$ is a (central) complex Wishart matrix with n degrees of freedom and covariance matrix $\boldsymbol{R}$ (denoted as $\boldsymbol{W} \sim \mathcal{W}_m(n, \boldsymbol{R})$), if the columns of the $m \times n$ matrix $\boldsymbol{H}$ are zero-mean independent complex Gaussian vectors with covariance matrix $\boldsymbol{R}$.

Lemma 2: The joint pdf of the eigenvalues of $\boldsymbol{W} \triangleq \boldsymbol{H}\boldsymbol{H}^\dagger \sim \mathcal{W}_m(n, \boldsymbol{R}_{m \times m})$ is identical to that of any $\boldsymbol{W}' \sim \mathcal{W}_{m'}(n, \mathrm{diag}(\mu_1, \ldots, \mu_{m'}))$ if $\mu_1 \geq \ldots \geq \mu_{m'} > \mu_{m'+1} = \ldots = \mu_m = 0$ are the eigenvalues of $\boldsymbol{R}_{m \times m}$.

Proof: Let $\boldsymbol{R} = \boldsymbol{Q}^\dagger \mathrm{diag}(\mu_1, \ldots, \mu_{m'}, 0, \ldots, 0)\boldsymbol{Q}$ be the eigenvalue decomposition of $\boldsymbol{R}$. Then, define $\sqrt{\boldsymbol{R}} \triangleq \boldsymbol{Q}^\dagger \mathrm{diag}(\sqrt{\mu_1}, \ldots, \sqrt{\mu_{m'}}, 0, \ldots, 0)\boldsymbol{Q}$ and $\boldsymbol{H}$ can be rewritten as $\boldsymbol{H} = \sqrt{\boldsymbol{R}}\boldsymbol{H}_0$ with $\boldsymbol{H}_0$ having i.i.d. $\mathcal{CN}(0,1)$ entries. We know that the eigenvalues of $\boldsymbol{H}\boldsymbol{H}^\dagger$ are identical to those of

$$
\begin{aligned}
\boldsymbol{H}^\dagger \boldsymbol{H} &= \boldsymbol{H}_0^\dagger \boldsymbol{R}\boldsymbol{H}_0 \\
&= (\boldsymbol{Q}\boldsymbol{H}_0)^\dagger \mathrm{diag}(\mu_1, \ldots, \mu_{m'}, 0, \ldots, 0)(\boldsymbol{Q}\boldsymbol{H}_0) \\
&= \widetilde{\boldsymbol{H}}_0^\dagger \mathrm{diag}(\mu_1, \ldots, \mu_{m'}, 0, \ldots, 0)\widetilde{\boldsymbol{H}}_0 \\
&= \widehat{\boldsymbol{H}}_0^\dagger \mathrm{diag}(\mu_1, \ldots, \mu_{m'})\widehat{\boldsymbol{H}}_0
\end{aligned}
$$

where $\widetilde{\boldsymbol{H}}_0 \triangleq \boldsymbol{Q}\boldsymbol{H}_0 \in \mathbb{C}^{m \times n}$ has i.i.d. entries as $\boldsymbol{H}_0$ does; $\widehat{\boldsymbol{H}}_0 \in \mathbb{C}^{m' \times n}$ is composed of the first m' rows of $\widetilde{\boldsymbol{H}}_0$ and its entries is thus i.i.d. as well. Finally, we prove the lemma using the fact that the eigenvalues of $\widehat{\boldsymbol{H}}_0^\dagger \mathrm{diag}(\mu_1, \ldots, \mu_{m'})\widehat{\boldsymbol{H}}_0$ are identical to those of

$$
\boldsymbol{W}' \triangleq (\mathrm{diag}(\sqrt{\mu_1}, \ldots, \sqrt{\mu_{m'}})\widehat{\boldsymbol{H}}_0)(\mathrm{diag}(\sqrt{\mu_1}, \ldots, \sqrt{\mu_{m'}})\widehat{\boldsymbol{H}}_0)^\dagger.
$$

$\blacksquare$

Lemma 3 ([19]–[22]): Let $\boldsymbol{W}$ be a central complex Wishart matrix $\boldsymbol{W} \sim \mathcal{W}_m(n, \boldsymbol{R})$, where the eigenvalues of $\boldsymbol{R}$ are distinct[6] and their ordered values are $\mu_1 > \ldots > \mu_m > 0$. Let $\lambda_1 > \ldots > \lambda_q > 0$ be the ordered positive eigenvalues of $\boldsymbol{W}$ with $q \triangleq \min\{m, n\}$. The joint pdf of $\boldsymbol{\lambda}$ conditionned on $\boldsymbol{\mu}$ is

$$
p(\boldsymbol{\lambda}|\boldsymbol{\mu}) = \begin{cases} K_{m,n}\mathrm{Det}(\boldsymbol{\Xi}_1) \prod_{i=1}^{m} \mu_i^{m-n-1} \lambda_i^{n-m} \prod_{i<j}^{m} \dfrac{\lambda_i - \lambda_j}{\mu_i - \mu_j}, & \text{if } n \geq m, \quad (19\text{a}) \\[2em] G_{m,n}\mathrm{Det}(\boldsymbol{\Xi}_2) \prod_{i<j}^{m} \dfrac{1}{(\mu_i - \mu_j)} \prod_{i<j}^{n} (\lambda_i - \lambda_j), & \text{if } n < m, \quad (19\text{b}) \end{cases}
$$

with $\boldsymbol{\Xi}_1 \triangleq \left[e^{-\lambda_j/\mu_i}\right]_{i,j=1}^{m}$ and

$$
\boldsymbol{\Xi}_2 \triangleq \begin{bmatrix} 1 & \mu_1 & \cdots & \mu_1^{m-n-1} & \mu_1^{m-n-1}e^{-\frac{\lambda_1}{\mu_1}} & \cdots & \mu_1^{m-n-1}e^{-\frac{\lambda_n}{\mu_1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_m & \cdots & \mu_m^{m-n-1} & \mu_m^{m-n-1}e^{-\frac{\lambda_1}{\mu_m}} & \cdots & \mu_m^{m-n-1}e^{-\frac{\lambda_n}{\mu_m}} \end{bmatrix}. \tag{20}
$$

$K_{m,n}$ and $G_{m,n}$ are normalization factors. In particular, for $\boldsymbol{R} = \mathbf{I}$, the joint pdf is

$$
P_{m,n}e^{-\sum_i \lambda_i} \prod_{i=1}^{q} \lambda_i^{|m-n|} \prod_{i<j}^{q} (\lambda_i - \lambda_j)^2. \tag{21}
$$

Now, let us define the eigen-exponents $\alpha_i \triangleq -\log \lambda_i / \log \mathsf{SNR}$, $i = 1, \ldots, q$, and $\beta_i \triangleq -\log \mu_i / \log \mathsf{SNR}$, $i = 1, \ldots, m$.

Lemma 4:

$$
\mathrm{Det}(\boldsymbol{\Xi}_1) \doteq \begin{cases} \mathsf{SNR}^{-E_{\boldsymbol{\Xi}_1}(\boldsymbol{\alpha},\boldsymbol{\beta})}, & \text{for } (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}^{(1)} \\[0.5em] \mathsf{SNR}^{-\infty}, & \text{otherwise}, \end{cases} \tag{22}
$$

where

$$
E_{\boldsymbol{\Xi}_1}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \triangleq \sum_{j=1}^{m} \sum_{i<j} (\alpha_i - \beta_j)^+, \tag{23}
$$

and

$$
\mathcal{R}^{(1)} \triangleq \left\{\alpha_1 \leq \ldots \leq \alpha_m, \ \beta_1 \leq \ldots \leq \beta_m, \ \text{and } \beta_i \leq \alpha_i, \ \text{for } i = 1, \ldots, m\right\}. \tag{24}
$$

Proof:

Please refer to [8] for details. ∎

---

[6]In the particular case where some eigenvalues of $\boldsymbol{R}$ are identical, we apply the l'Hospital rule to the pdf obtained, as shown in [21].

Lemma 5:

$$\text{Det}\,(\boldsymbol{\Xi}_2) \doteq \begin{cases} \text{SNR}^{-E_{\boldsymbol{\Xi}_2}(\boldsymbol{\alpha},\boldsymbol{\beta})}, & \text{for } (\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{R}^{(2)} \\ \\ \text{SNR}^{-\infty}, & \text{otherwise,} \end{cases} \tag{25}$$

where

$$E_{\boldsymbol{\Xi}_2}(\boldsymbol{\alpha},\boldsymbol{\beta}) \triangleq \sum_{i=1}^{n}(m-n-1)\beta_i + \sum_{i=n+1}^{m}(m-i)\beta_i + \sum_{j=1}^{n}\sum_{i<j}(\alpha_i-\beta_j)^+ + \sum_{j=n+1}^{m}\sum_{i=1}^{n}(\alpha_i-\beta_j)^+ \tag{26}$$

and

$$\mathcal{R}^{(2)} \triangleq \{\alpha_1 \leq \ldots \leq \alpha_n,\ \beta_1 \leq \ldots \leq \beta_m,\ \text{and } \beta_i \leq \alpha_i,\ \text{for } i=1,\ldots,n\}. \tag{27}$$

Proof: First, we have

$$\text{Det}(\boldsymbol{\Xi}_2) = \prod_{i=1}^{m} \mu_i^{m-n-1} \text{Det} \begin{bmatrix} \mu_1^{-(m-n-1)} & \cdots & 1 & e^{-\lambda_1/\mu_1} & \cdots & e^{-\lambda_n/\mu_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mu_m^{-(m-n-1)} & \cdots & 1 & e^{-\lambda_1/\mu_m} & \cdots & e^{-\lambda_n/\mu_m} \end{bmatrix}. \tag{28}$$

Then, let us denote the determinant in the right hand side (RHS) of (28) as D and we rewrite it as

$$D = \text{Det} \begin{bmatrix} d_{1,m}^{(m-n-1)} & \cdots & 0 & e^{-\lambda_1/\mu_1} - e^{-\lambda_1/\mu_m} & \cdots & e^{-\lambda_n/\mu_1} - e^{-\lambda_n/\mu_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ d_{m-1,m}^{(m-n-1)} & \cdots & 0 & e^{-\lambda_1/\mu_{m-1}} - e^{-\lambda_1/\mu_m} & \cdots & e^{-\lambda_n/\mu_{m-1}} - e^{-\lambda_n/\mu_m} \\ \mu_m^{-(m-n-1)} & \cdots & 1 & e^{-\lambda_1/\mu_m} & \cdots & e^{-\lambda_n/\mu_m} \end{bmatrix} \tag{29}$$

$$\doteq \text{Det} \begin{bmatrix} d_{1,m}^{(m-n-1)} & \cdots & d_{1,m}^{(1)} & e^{-\lambda_1/\mu_1} & \cdots & e^{-\lambda_n/\mu_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ d_{m-1,m}^{(m-n-1)} & \cdots & d_{m-1,m}^{(1)} & e^{-\lambda_1/\mu_{m-1}} & \cdots & e^{-\lambda_n/\mu_{m-1}} \end{bmatrix} \prod_{i=1}^{n}\left(1 - e^{-\lambda_i/\mu_m}\right) \tag{30}$$

where $d_{i,j}^{(k)} \triangleq \mu_i^{-k} - \mu_j^{-k}$ and the product term in (30) is obtained since $1 - e^{-(\lambda_i/\mu_m - \lambda_i/\mu_j)} \doteq 1 - e^{-\lambda_i/\mu_m}$ for all $j < m$. Let us denote the determinant in (30) as D_m. Then, by multiplying the first column in D_m with μ_m^{m-n-1} and noting that $\mu_m^{m-n-1}d_{i,m}^{(m-n-1)} = 1 - (\mu_m/\mu_i)^{m-n-1} \approx 1$, the first column of D_m becomes all 1. Now, by eliminating the first $m-2$ "1"s of the first column by subtracting all rows by the last row as in (29) and (30), we have $\mu_m^{m-n-1}D_m \doteq$

$\prod_{i=1}^{n} \left(1 - e^{-\lambda_i/\mu_m}\right) D_{m-1}$. By continuing reducing the dimension, we get

$$\text{Det}(\boldsymbol{\Xi}_2) \doteq \text{Det}\left[e^{-\lambda_j/\mu_i}\right]_{i,j=1}^{n} \prod_{i=1}^{n+1} \mu_i^{m-n-1} \prod_{i=n+2}^{m} \mu_i^{m-i}$$

$$\cdot \prod_{i=1}^{n} \prod_{j=n+1}^{m} \left(1 - e^{-\lambda_i/\mu_j}\right)$$

from which we prove the lemma, by applying (22). ∎

With the two preceding lemmas, we have the following lemma that provides the asymptotical pdf of $\boldsymbol{\alpha}$ conditionned on $\boldsymbol{\beta}$ in the high SNR regime.

Lemma 6:

$$p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \doteq \begin{cases} \mathsf{SNR}^{-E(\boldsymbol{\alpha}|\boldsymbol{\beta})}, & \text{for } (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}_{\boldsymbol{\alpha}|\boldsymbol{\beta}}, \\ \mathsf{SNR}^{-\infty}, & \text{otherwise,} \end{cases} \tag{31}$$

where

$$E(\boldsymbol{\alpha}|\boldsymbol{\beta}) \triangleq \sum_{i=1}^{q}(n+1-i)\alpha_i + \sum_{i=1}^{q}(i-n-1)\beta_i + \sum_{j=1}^{q}\sum_{i<j}(\alpha_i - \beta_j)^+ + \sum_{j=q+1}^{m}\sum_{i=1}^{q}(\alpha_i - \beta_j)^+, \tag{32}$$

and

$$\mathcal{R}_{\boldsymbol{\alpha}|\boldsymbol{\beta}} \triangleq \{\alpha_1 \leq \ldots \leq \alpha_q, \ \beta_1 \leq \ldots \leq \beta_m, \ \text{and } \beta_i \leq \alpha_i, \ \text{for } i = 1, \ldots, q\}. \tag{33}$$

Proof: For $n \geq m$, applying the variable changes to (19a), we have

$$p(\boldsymbol{\alpha}|\boldsymbol{\beta}) = K_{m,n}(\log \mathsf{SNR})^l \prod_{i=1}^{m} \mathsf{SNR}^{-(n-m+1)\alpha_i} \mathsf{SNR}^{-(m-n-1)\beta_i}$$

$$\cdot \prod_{j=1}^{m} \prod_{i<j} (\mathsf{SNR}^{-\alpha_i} - \mathsf{SNR}^{-\alpha_j})(\mathsf{SNR}^{-\beta_i} - \mathsf{SNR}^{-\beta_j})^{-1}$$

$$\cdot \text{Det}\left[\exp\left(-\mathsf{SNR}^{-(\alpha_j - \beta_i)}\right)\right].$$

The high SNR exponent of the quantity $\text{Det}\left[\exp\left(-\mathsf{SNR}^{-(\alpha_j - \beta_i)}\right)\right]$ is calculated in Lemma 4. From (22), we only need to consider $\alpha_i \geq \beta_i, \forall i$, so that $p(\boldsymbol{\alpha}|\boldsymbol{\beta})$ does not decay exponentially. Therefore, we have

$$p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \doteq \mathsf{SNR}^{-\left(\sum_{i=1}^{m}(n+1-i)\alpha_i + \sum_{i=1}^{m}(i-n-1)\beta_i + \sum_{j=1}^{m}\sum_{i<j}(\alpha_i - \beta_j)^+\right)}, \tag{34}$$

if $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}^{(1)}$ and $p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \doteq \mathsf{SNR}^{-\infty}$ otherwise.

For $n < m$, with (19b) and (25), we get

$$p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \doteq \prod_{i=1}^{n} \mathsf{SNR}^{-(m-n-1)\beta_i} \prod_{i=n+1}^{m} \mathsf{SNR}^{-(m-i)\beta_i}$$

$$\cdot \prod_{j=1}^{n} \prod_{i<j} \mathsf{SNR}^{-(\alpha_i-\beta_j)^+} \prod_{j=n+1}^{m} \prod_{i=1}^{n} \mathsf{SNR}^{-(\alpha_i-\beta_j)^+} \qquad (35)$$

$$\cdot \prod_{i=1}^{n} \mathsf{SNR}^{-(n+1-i)\alpha_i} \prod_{i=1}^{m} \mathsf{SNR}^{(m-i)\beta_i}.$$

for $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}^{(2)}$ and $p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \doteq \mathsf{SNR}^{-\infty}$ otherwise. Combining the two cases, we prove the lemma. ∎

When $\boldsymbol{R} = \mathbf{I}$, *i.e.*, $\mu_1 = \ldots = \mu_m = 1$, the joint pdf of $\boldsymbol{\alpha}$ is found in [11] as shown in the following lemma.

Lemma 7:

$$p(\boldsymbol{\alpha}) \doteq \begin{cases} \mathsf{SNR}^{-\sum_{i=1}^{q}(m+n+1-2i)\alpha_i}, & \text{for } \boldsymbol{\alpha} \in \mathcal{R}_{\boldsymbol{\alpha}}, \\ \mathsf{SNR}^{-\infty}, & \text{otherwise}, \end{cases} \qquad (36)$$

with $\mathcal{R}_{\boldsymbol{\alpha}} \triangleq \{0 \leq \alpha_1 \leq \ldots \leq \alpha_q\}$.

This lemma can be justified either by using (21) or by setting $\beta_i = 0$, $\forall i$ in (32).

Lemma 8 ([23]): Let $\boldsymbol{M}$ be any $m \times n$ random matrix and $\boldsymbol{T}$ be any $m \times m$ non-singular matrix whose singular values satisfy $\sigma_{\min}(\boldsymbol{T}) \doteq \sigma_{\max}(\boldsymbol{T}) \doteq \mathsf{SNR}^0$. Define $q \triangleq \min\{m, n\}$ and $\tilde{\boldsymbol{M}} \triangleq \boldsymbol{T}\boldsymbol{M}$. Let $\sigma_1(\boldsymbol{M}) \geq \ldots \geq \sigma_q(\boldsymbol{M})$ and $\sigma_1(\tilde{\boldsymbol{M}}) \geq \ldots \geq \sigma_q(\tilde{\boldsymbol{M}})$ be the ordered singular values of $\boldsymbol{M}$ and $\tilde{\boldsymbol{M}}$, Then, we have

$$\sigma_i(\tilde{\boldsymbol{M}}) \doteq \sigma_i(\boldsymbol{M}), \quad \forall i.$$

APPENDIX II

PROOF OF THEOREM 1

The following lemma will be used repeatedly in the most of the proofs.

Lemma 9: Let $\mathcal{I}_k \triangleq [\, p_k, p_{k-1} \,]$, $k = 1, \ldots, N$, be N consecutively joint intervals with $p_N \triangleq -\infty$, $p_0 \triangleq \tilde{n}_0$, and p_k's are defined as in (15). Then, we have

$$c_i = 1 - i + \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - i}{k} \right\rfloor, \quad \text{for } i \in \mathcal{I}_k. \qquad (37)$$

Proof: c_i defined by (7) is the minimum of N sequences corresponding to the N values of k. It is enough to show that each of the N sequences dominates in a consecutive manner. We omit the details here. ∎

A. Sketch of the Proof

The proof will be by induction on N. From lemma 7, the theorem is trivial for $N = 1$. Suppose the theorem holds for some N and $\mathbf{\Pi} \triangleq \boldsymbol{H}_1 \cdots \boldsymbol{H}_N$, we would like to show that it is also true for $N + 1$ and $\mathbf{\Pi}' \triangleq \boldsymbol{H}_1 \cdots \boldsymbol{H}_{N+1}$. For simplicity, the "primed" notations (*e.g.*, $\boldsymbol{\alpha}'$, $\boldsymbol{n}'$, $\tilde{\boldsymbol{n}}'$, $\boldsymbol{c}'$, $n'_{\min}$, etc.) will be used for the respective parameters of $\mathbf{\Pi}'$. Note that $\mathbf{\Pi}'(\mathbf{\Pi}')^\dagger \sim \mathcal{W}_{n_0}(n_{N+1}, \mathbf{\Pi}\mathbf{\Pi}^\dagger)$ for a given $\mathbf{\Pi}$, since $\mathbf{\Pi}' = \mathbf{\Pi}\boldsymbol{H}_{N+1}$. According to lemma 2, the pdf of the eigenvalues $\boldsymbol{\lambda}'$ of $\mathbf{\Pi}'(\mathbf{\Pi}')^\dagger$ is identical to that of $\mathcal{W}_{n_{\min}}(n_{N+1}, \operatorname{diag}(\boldsymbol{\lambda}))$. Hence, the pdf of $\boldsymbol{\alpha}'$ can be obtained as the marginal pdf of $(\boldsymbol{\alpha}', \boldsymbol{\alpha})$

$$
p(\boldsymbol{\alpha}') = \int_{\mathbb{R}^{n_{\min}}} p(\boldsymbol{\alpha}', \boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha}
$$

$$
= \int_{\mathbb{R}^{n_{\min}}} p(\boldsymbol{\alpha}'|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha}
$$

$$
\doteq \int_{\mathcal{R}} \mathsf{SNR}^{-E(\boldsymbol{\alpha}'|\boldsymbol{\alpha})} \mathsf{SNR}^{-E(\boldsymbol{\alpha})} \mathrm{d}\boldsymbol{\alpha} \tag{38}
$$

$$
\doteq \mathsf{SNR}^{-\hat{E}(\boldsymbol{\alpha}')} \tag{39}
$$

where (38) comes from lemma 6 and our assumption that (5) holds for N, with

$$
\mathcal{R} \triangleq \mathcal{R}_{\boldsymbol{\alpha}'|\boldsymbol{\alpha}} \cap \mathcal{R}_{\boldsymbol{\alpha}}
$$
$$
= \left\{ 0 \le \alpha'_1 \le \ldots \le \alpha'_{n'_{\min}}, \ 0 \le \alpha_1 \le \ldots \le \alpha_{n_{\min}}, \text{ and } \alpha_i \le \alpha'_i, \text{ for } i = 1, \ldots, n'_{\min} \right\} \tag{40}
$$

being the feasible region; the exponent $\hat{E}(\boldsymbol{\alpha}')$ in (39) is defined by

$$
\hat{E}(\boldsymbol{\alpha}') = \min_{\boldsymbol{\alpha} \in \mathcal{R}} E(\boldsymbol{\alpha}', \boldsymbol{\alpha}) \tag{41}
$$

with $E(\boldsymbol{\alpha}', \boldsymbol{\alpha}) \triangleq E(\boldsymbol{\alpha}'|\boldsymbol{\alpha}) + E(\boldsymbol{\alpha})$. From (32) and (6),

$$
E(\boldsymbol{\alpha}', \boldsymbol{\alpha}) = \sum_{i=1}^{n'_{\min}} (n_{N+1} - i + 1)\alpha'_i + \sum_{j=1}^{n'_{\min}} \left((j - 1 - n_{N+1} + c_j)\alpha_j + \sum_{i<j} (\alpha'_i - \alpha_j)^+ \right)
$$
$$
+ \sum_{j=n'_{\min}+1}^{n_{\min}} \left(c_j \alpha_j + \sum_{i=1}^{n'_{\min}} (\alpha'_i - \alpha_j)^+ \right). \tag{42}
$$

It remains to show $\hat{E}(\boldsymbol{\alpha}') = E'(\boldsymbol{\alpha}') \triangleq \sum_i c_i \alpha'_i$ with

$$
c'_i \triangleq 1 - i + \min_{k=1,\ldots,N+1} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}'_l - i}{k} \right\rfloor, \quad i = 1, \ldots, n'_{\min} \tag{43}
$$

by solving the optimization problem (41), which is accomplished in the rest of the section.
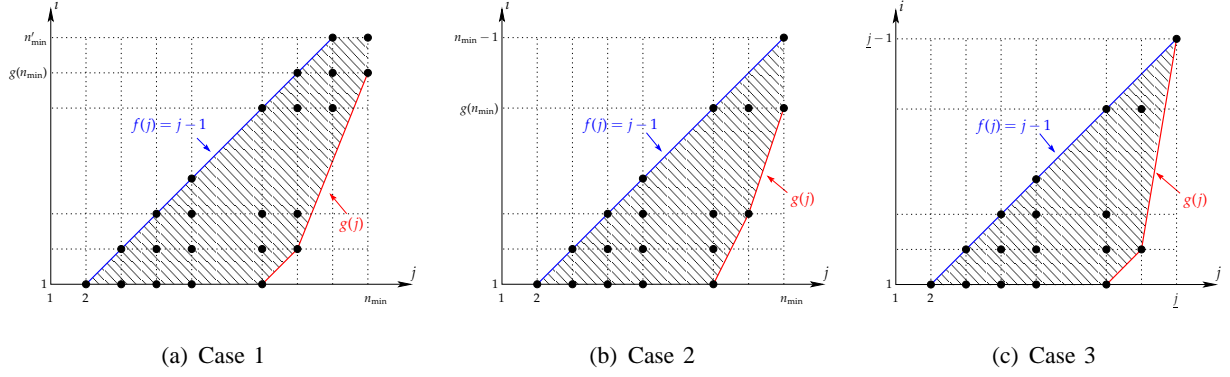
(a) Case 1   (b) Case 2   (c) Case 3

Fig. 6. For each j, the black dots represent the α''s that are freed by α_j. Therefore, we can get the total number of freed α'_i by counting the black dots in row i. More precisely, there are $\lfloor g^{-1}(i) \rfloor - \lceil f^{-1}(i) \rceil + 1 = \lfloor g^{-1}(i) \rfloor - i$ black dots for $i \leq g(n_{\min})$, and $n_{\min} - \lceil f^{-1}(i) \rceil + 1 = n_{\min} - i$ black dots for $i > g(n_{\min})$.

B. Solving the Optimization Problem

1) Case 1 [$n_{N+1} < \tilde{n}_0$]: In this case, we have $n'_{\min} = \tilde{n}'_0 = n_{N+1}$. Minimization of $E(\boldsymbol{\alpha}, \boldsymbol{\alpha}')$ of (42) with respect to (*w.r.t.*) $\boldsymbol{\alpha}$ can be decomposed into $n_{\min}$ minimizations *w.r.t.* $\alpha_1, \ldots, \alpha_{n_{\min}}$ successively, *i.e.*,

$$\min_{\boldsymbol{\alpha}} = \min_{\alpha_{n_{\min}}} \cdots \min_{\alpha_1}.$$

We start with α_1. From (33), the feasible region of α_1 is $0 \leq \alpha_1 \leq \alpha'_1$. Since the only α_1-related term in (42) is $(c_1 - n_{N+1})\alpha_1$ and $c_1 - n_{N+1} > 0$ for $n_{N+1} < \tilde{n}_0$, we have $\alpha_1^* = 0$. Now, suppose that the minimization *w.r.t.* $\alpha_1, \ldots, \alpha_{j-1}$ is done and that we would like to minimize *w.r.t.* α_j. For α_j, $j \leq n'_{\min}$, we set the initial region as

$$0 \leq \alpha'_1 \leq \cdots \leq \alpha'_{j-1} \leq \alpha_j \leq \alpha'_j$$

in which we have $\sum_{i<j} (\alpha'_i - \alpha_j)^+ = 0$. The feasibility conditions in (40) require that α_j must not go right across α'_j. The only choice is therefore to go to the left. Each time α_j goes across an α'_i from the right to the left, $(\alpha'_i - \alpha_j)^+$ increases by $\alpha'_i - \alpha_j$, which increases the coefficient of α'_i by 1 and decreases the coefficient of α_j by 1. It can be shown that, to minimize the value of $E(\boldsymbol{\alpha}, \boldsymbol{\alpha}')$ *w.r.t.* α_j, α_j is allowed to cross α'_i only when the current coefficient of α_j in (42) is positive[7]. So, α_j stops moving only in the following two cases : 1) it hits the left extreme, 0; and 2) its coefficient achieves 0 when it is in the interval $[\alpha'_k, \alpha'_{k+1}]$ for some

---

[7]When the coefficient of α_i in (42) is positive, decreasing α_i decreases $E(\boldsymbol{\alpha}, \boldsymbol{\alpha}')$.

$k < j$. Either case, α_j-related terms are gone and what remain are the α_i''s "freed" by α_j from $\sum_{i<j} (\alpha_i' - \alpha_j)^+$. Same reasoning applies to α_j for $j > n'_{\min}$, except that the initial region is set to $0 \leq \alpha_1' \leq \cdots \leq \alpha'_{n'_{\min}} \leq \alpha_j$.

Therefore, the optimization problem can be solved by counting the total number of freed α_i''s. As shown in Fig. 6(a), when j is small, the initial coefficient of α_j is large and thus α_j can free out $\alpha_{j-1}', \ldots, \alpha_1'$. We have $\alpha_j^* = 0$, which corresponds to the first stopping condition. For large j, the initial coefficient of α_j is not large enough and only $\alpha_{j-1}', \ldots, \alpha'_{g(j)}$ is freed, which corresponds to the second stopping condition. With the above reasoning, we can get $g(j)$

$$g(j) = \begin{cases} j - 1 - (j - 1 - n_{N+1} + c_j) + 1, & \text{for } j \leq n'_{\min}, \\ n_{N+1} - c_j + 1, & \text{for } j > n'_{\min}. \end{cases} \tag{44}$$

From (44) and (7), we get

$$g(j) = n_{N+1} - \min_{k=1,\ldots,N} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - (k+1)j}{k} \right\rfloor, \tag{45}$$

and

$$\lfloor g^{-1}(i) \rfloor = \min_{k=1,\ldots,N} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - k(n_{N+1} - i)}{k+1} \right\rfloor. \tag{46}$$

Now, $\hat{E}(\boldsymbol{\alpha}')$ can be obtained[8] from Fig. 6(a)

$$\hat{E}(\boldsymbol{\alpha}') = \sum_{i=1}^{n'_{\min}} (n_{N+1} - i + 1)\alpha_i' + \sum_{i=1}^{g(n_{\min})} (\lfloor g^{-1}(i) \rfloor - i)\alpha_i' + \sum_{i=g(n_{\min})+1}^{n'_{\min}} (n_{\min} - i)\alpha_i'$$

$$= \sum_{i=1}^{g(n_{\min})} \left(1 - 2i + n_{N+1} + \lfloor g^{-1}(i) \rfloor\right)\alpha_i' + \sum_{i=g(n_{\min})+1}^{n'_{\min}} \left(1 - 2i + n_{N+1} + n_{\min}\right)\alpha_i'$$

$$= \sum_{i=1}^{g(n_{\min})} \left(1 - i + \min_{k=2,\ldots,N+1} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l' - i}{k} \right\rfloor\right)\alpha_i' + \sum_{i=g(n_{\min})+1}^{n'_{\min}} \left(1 - 2i + n_{N+1} + n_{\min}\right)\alpha_i'$$

$$\tag{47}$$

$$= \sum_{i=1}^{n'_{\min}} \left(1 - i + \min_{k=1,\ldots,N+1} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l' - i}{k} \right\rfloor\right)\alpha_i' \tag{48}$$

$$= E'(\boldsymbol{\alpha}'), \tag{49}$$

---

[8]In the above minimization procedure, we ignored the feasibility condition $\alpha_j \geq \alpha_k, \ \forall j > k$. A more careful analysis can reveal that it is always satisfied with the described procedure.

where (47) is from (46) and the fact that $\tilde{n}_0' = n_{N+1}$, $\tilde{n}_l' = \tilde{n}_{l-1}$, $l = 1, \ldots, N+1$; (48) can be derived from lemma 9, since $p_1' = n_{N+1} + \tilde{n}_0 - \tilde{n}_1 = g(n_{\min})$ and therefore the term $\min_k$ in (48) is dominated by $k \geq 2$ for $i \leq g(n_{\min})$ and by $k = 1$ for $i > g(n_{\min})$, corresponding to the two terms in (47), respectively.

2) Case 2 [$n_{N+1} \in [\tilde{n}_0, \tilde{n}_1)$]: In this case, we have $n_{\min}' = n_{\min}$ and $\tilde{n}_1' = n_{N+1}$. From (42),

$$E(\boldsymbol{\alpha}', \boldsymbol{\alpha}) = \sum_{i=1}^{n_{\min}'} (n_{N+1} - i + 1)\alpha_i' + \sum_{j=1}^{n_{\min}'} \left((j - 1 - n_{N+1} + c_j)\alpha_j + \sum_{i<j} (\alpha_i' - \alpha_j)^+ \right). \quad (50)$$

Since $j - 1 - n_{N+1} + c_j > 0$, $\forall j \leq n_{\min}'$, the minimization of $E(\boldsymbol{\alpha}', \boldsymbol{\alpha})$ w.r.t. $\boldsymbol{\alpha}$ is in exactly the same manner as in the previous case. Therefore, $\hat{E}(\boldsymbol{\alpha}')$ can be obtained from Fig. 6(b) with $g(j)$ in the same form as (45)

$$\hat{E}(\boldsymbol{\alpha}') = \sum_{i=1}^{n_{\min}'} (n_{N+1} - i + 1)\alpha_i' + \sum_{i=1}^{g(n_{\min})} (\lfloor g^{-1}(i) \rfloor - i)\alpha_i' + \sum_{i=g(n_{\min})+1}^{n_{\min}'} (n_{\min} - i)\alpha_i'$$

$$= E'(\boldsymbol{\alpha}'). \quad (51)$$

3) Case 3 [$n_{N+1} \in [\tilde{n}_1, \infty)$]: As in the last case, we have $n_{\min}' = n_{\min}$ and the same $E(\boldsymbol{\alpha}', \boldsymbol{\alpha})$ as defined in (50). Without loss of generality, we assume that $n_{N+1} \in [\tilde{n}_{k^*}, \tilde{n}_{k^*+1})$ for some $k^* \in [1, N]$ (we set $\tilde{n}_{N+1} \triangleq \infty$). Then, we have

$$\tilde{n}_l' = \tilde{n}_l, \quad \text{for } l = 1, \ldots, k^*, \quad (52)$$

and

$$p_{k^*} < p_{k^*}' \leq p_{k^*-1} = p_{k^*-1}' \leq \cdots \leq p_1 = p_1'. \quad (53)$$

Unlike the previous case, $j - 1 - n_{N+1} + c_j$ is not always positive. Let $\underline{j}$ be the smallest integer such that the coefficient $j - 1 - n_{N+1} + c_j$ of α_j in (50) is zero. It is obvious that for $j \geq \underline{j}$, $\alpha_j^* = \alpha_j'$. Hence, we have

$$\hat{E}(\boldsymbol{\alpha}') = \sum_{i=1}^{n_{\min}'} (n_{N+1} - i + 1)\alpha_i' + \sum_{i=1}^{\underline{j}-1} (\lfloor g^{-1}(i) \rfloor - i)\alpha_i' + \sum_{j=\underline{j}}^{n_{\min}'} (j - 1 - n_{N+1} + c_j)\alpha_j'$$

where the second term is from Fig. 6(c). Furthermore, we can show that $\underline{j} \leq p_{k^*}'$, since $p_{k^*}' - 1 - n_{N+1} + c_{p_{k^*}'} = 0$. Therfore, we get

$$\hat{E}(\boldsymbol{\alpha}') = \sum_{i=1}^{\underline{j}-1} \left(1 - 2i + n_{N+1} + \lfloor g^{-1}(i) \rfloor \right) \alpha_i' + \sum_{i=\underline{j}}^{p_{k^*}'-1} (n_{N+1} - i + 1)\alpha_i' + \sum_{i=p_{k^*}'}^{n_{\min}'} c_i \alpha_i'. \quad (54)$$

Now, we would like to show that the coefficient of α'_i in (54) coincides with c'_i. First, for $i \leq \underline{j}-1$, $i \in \mathcal{I}'_{k^*+1} \cup \cdots \cup \mathcal{I}'_N$ and lemma 9 implies that

$$
1 - 2i + n_{N+1} + \lfloor g^{-1}(i) \rfloor = 1 - i + \min_{k=2,\ldots,N+1} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}'_l - i}{k} \right\rfloor
$$

$$
= 1 - i + \min_{k=1,\ldots,N+1} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}'_l - i}{k} \right\rfloor
$$

$$
= c'_i.
$$

Then, for $i \geq p'_{k^*}$, we have

$$
i \in \left(\mathcal{I}'_{k^*} \cup \cdots \cup \mathcal{I}'_1 \right) \cap \left(\mathcal{I}_{k^*} \cup \cdots \cup \mathcal{I}_1 \right).
$$

Hence,

$$
c'_i = 1 - i + \min_{k=1,\ldots,k^*} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}'_l - i}{k} \right\rfloor
$$

$$
= 1 - i + \min_{k=1,\ldots,k^*} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - i}{k} \right\rfloor \tag{55}
$$

$$
= c_i,
$$

where (55) is from (52) and (53). Finally, for $i \in [\underline{j}, p'_{k^*})$, let us rewrite $i = p'_{k^*} - \Delta_i$. Since $i - 1 - n_{N+1} + c_i = 0, \forall i \in [\underline{j}, p'_{k^*})$, we have

$$
\left\lfloor \frac{\sum_{l=0}^{k^*} \tilde{n}_l - i - k^* n_{N+1}}{k^*} \right\rfloor = \left\lfloor \frac{\sum_{l=0}^{k^*} \tilde{n}_l - p'_{k^*} + \Delta_i - k^* n_{N+1}}{k^*} \right\rfloor
$$

$$
= \left\lfloor \frac{\Delta_i}{k^*} \right\rfloor
$$

$$
= 0,
$$

from which we have $\Delta_i \in [0, k^* - 1]$ and

$$
c'_i = \left\lfloor \frac{\sum_{l=0}^{k^*} \tilde{n}_l + n_{N+1} - i}{k^* + 1} \right\rfloor + 1 - i
$$

$$
= \left\lfloor \frac{\sum_{l=0}^{k^*} \tilde{n}_l + n_{N+1} - p'_{k^*} + \Delta_i}{k^* + 1} \right\rfloor + 1 - i
$$

$$
= 1 + n_{N+1} - i.
$$

The proof is complete.

C. *Proof of Theorem 6*

To prove the first case, we use induction on N. Suppose that it is true for N, which means that the joint pdf of $\boldsymbol{\alpha}(\boldsymbol{\Pi}_g \boldsymbol{\Pi}_g^\dagger)$ is the same as that of $\boldsymbol{\alpha}(\boldsymbol{\Pi}\boldsymbol{\Pi}^\dagger)$. Furthermore, we know by lemma 8 that $\boldsymbol{\alpha}(\boldsymbol{\Pi}_g \boldsymbol{T}_{N,N+1} \boldsymbol{T}_{N,N+1}^\dagger \boldsymbol{\Pi}_g^\dagger) = \boldsymbol{\alpha}(\boldsymbol{\Pi}_g \boldsymbol{\Pi}_g^\dagger)$. Same steps as (38)(39) complete the proof.

To prove the second statement, we perform a singular value decomposition on the matrices $\boldsymbol{T}_{i,i+1}$'s and then apply the first statement.

APPENDIX III

PROOF OF THEOREM 2 AND THEOREM 3

A. *Proof of Theorem 2*

Let

$$c_i^{(m)} \triangleq 1 - i + \min_{k=1,\ldots,m} \left\lfloor \frac{\sum_{l=0}^{k} \tilde{n}_l - i}{k} \right\rfloor, \quad i = 1, \ldots, n_{\min}.$$

What we should prove is that

$$c_i^{(N)} = c_i^{(k)}, \quad \text{for } i = 1, \ldots, n_{\min}$$

if and only if (8) is true. To this end, it is enough to show that

$$c_i^{(N)} = c_i^{(N-1)} \quad \text{for } i = 1, \ldots, n_{\min} \tag{56}$$

if and only if $p_{N-1} \leq N - 1$, that is, $(N-1)(\tilde{n}_N + 1) \geq \sum_{l=0}^{N-1} \tilde{n}_l$, and then apply the result successively to show the theorem.

1) The Direct Part: The direct part is to show that, if $p_{N-1} \leq N-1$, then (56) is true. From lemma 9, we see that $c_i^{(N)} = c_i^{(N-1)}$, $\forall i \geq p_{N-1}$. Hence, when $p_{N-1} \leq 1$, (56) holds. Now, let us consider the case $p_{N-1} > 1$. We would like to show that $c_i^{(N)} = c_i^{(N-1)}$ for $i \in [1, p_{N-1}]$. Let $j \triangleq p_{N-1} - i \in [0, p_{N-1} - 1]$. Then, we rewrite the two quantities

$$\left\lfloor \frac{\sum_{l=0}^{N} \tilde{n}_l - i}{N} \right\rfloor = \tilde{n}_N + \left\lfloor \frac{j}{N} \right\rfloor \tag{57}$$

$$\left\lfloor \frac{\sum_{l=0}^{N-1} \tilde{n}_l - i}{N - 1} \right\rfloor = \tilde{n}_N + \left\lfloor \frac{j}{N - 1} \right\rfloor \tag{58}$$

that are identical for $p_{N-1} \leq N - 1$, which proves that $c_i^{(N)} = c_i^{(N-1)}$. The proof for the direct part is complete.

2) Converse: If $p_{N-1} > N - 1$, then from (57) and (58), we have $c_i^{(N)} \neq c_i^{(N-1)}$ at least for $j = N - 1$, that is, $i = p_{N-1} - (N - 1)$. The proof is complete.

B. Proof of Theorem 3

The direct part of the theorem is trivial. To show the converse, let $\tilde{\boldsymbol{n}} \triangleq (\tilde{n}_0, \tilde{n}_1, \ldots, \tilde{n}_N)$ and $\tilde{\boldsymbol{n}}' \triangleq (\tilde{n}'_0, \tilde{n}'_1, \ldots, \tilde{n}'_{N'})$ be the two concerned minimal forms. In addition, we assume, without loss of generality, that

$$\tilde{n}_1 = \cdots = \tilde{n}_{i_1}, \ldots, \tilde{n}_{i_{M-1}+1} = \cdots = \tilde{n}_{i_M}$$

$$\tilde{n}'_1 = \cdots = \tilde{n}'_{i'_1}, \ldots, \tilde{n}'_{i'_{M'-1}+1} = \cdots = \tilde{n}'_{i'_{M'}}$$

with $i_M \leq N$ and $i'_{M'} \leq N'$. Now, let us define $c_{0i} \triangleq c_i - (1 - i)$ with c_i defined in (37). It can be shown that M intervals are non-trivial with $|\mathcal{I}_{i_k}| \neq 0$, $k = 1, \ldots, M$. The values of c_{0i}'s are in the following form

$$\ldots, \overbrace{\underbrace{\tilde{n}_{i_M}, \ldots, \tilde{n}_{i_M}}_{i_M}}^{|\mathcal{I}_{i_M}|}, \overbrace{\underbrace{\tilde{n}_{i_M} - 1, \ldots, \tilde{n}_{i_M} - 1}_{i_{M-1}}, \ldots, \underbrace{\tilde{n}_{i_{M-1}}, \ldots, \tilde{n}_{i_{M-1}}}_{i_{M-1}}}^{|\mathcal{I}_{i_{M-1}}|}, \ldots, \overbrace{\tilde{n}_2 - 1, \ldots, \tilde{n}_1 + 1, \tilde{n}_1}^{|\mathcal{I}_1|}.$$

Same arguments also apply to $\tilde{\boldsymbol{n}}$ with M' and i', etc. It is then not difficult to see that to have exactly the same c_{0i}'s (thus, same c_i's), we must have $N = N'$ and

$$\tilde{n}_i = \tilde{n}'_i, \ \forall i = 0, \ldots, N,$$

that is, the same minimal form.

APPENDIX IV

PROOF OF THEOREM 5

A. Sketch of the Proof

To prove the theorem, we will first show the following equivalence relations :

$$(R_1^{(N)}(k), R_3^{(N)}(i, k)) \overset{(a)}{\Longleftrightarrow} (R_1^{(N)}(k), R_2^{(N)}(i)), \quad \forall i, k;$$

$$R_3^{(N)}(i, k) \overset{(b)}{\Longleftrightarrow} R_3^{(N)}(N - 1, k), \quad \forall i, k;$$

$$(R_1^{(N)}(k), R_2^{(N)}(N - 1)) \overset{(c)}{\Longleftrightarrow} (R_1^{(N)}(k), R_2^{(N)}(i) \text{ with ordered } \boldsymbol{n});$$

$$(R_1^{(N)}(k), R_2^{(N)}(i) \text{ with ordered } \boldsymbol{n}) \overset{(d)}{\Longleftrightarrow} (R_1^{(N)}(k), R_2^{(N)}(N - 1) \text{ with ordered and minimal } \boldsymbol{n}).$$

1) Equivalences (a) *and* (b)*:* The direct parts of (a), (b), and (d) are immediate since the RHS are particular cases of the left hand side (LHS). To show the reverse part of (a), we rewrite

$$d_{(n_0,...,n_N)}(k) = d_{(n_0-k,...,n_N-k)}(0) \tag{59}$$

$$= \min_{j \geq 0} d_{(n_0-k,...,n_i-k)}(j) + d_{(j,n_{i+1}-k,...,n_N-k)}(0) \tag{60}$$

$$= \min_{j' \geq k} d_{(n_0,...,n_i)}(j') + d_{(j',n_{i+1},...,n_N)}(k) \tag{61}$$

where R_1 is used twice in (59) and (61); R_2 is used in (60). As for (b), if $R_3^{(N)}(N-1,k)$ holds, then

$$d_{(n_0,...,n_N)}(k) = \min_{j \geq k} d_{(n_0,...,n_{N-1})}(j) + d_{(j,n_N)}(k) \tag{62}$$

$$= \min_{j' \geq j \geq k} d_{(n_0,...,n_{N-2})}(j') + d_{(j',n_{N-1})}(j) + d_{(j,n_N)}(k) \tag{63}$$

$$= \min_{j' \geq k} d_{(n_0,...,n_{N-2})}(j') + d_{(j',n_{N-1},n_N)}(k) \tag{64}$$

which proves $R_3^{(N)}(N-2,k)$. By continuing the process, we can show that $R_3^{(N)}(i,k)$ is true for all i, provided $R_3^{(N)}(N-1,k)$ holds.

2) Equivalences (c) *and* (d)*:* Through (a) and (b), one can verify that the LHS of (c) is equivalent to the RHS of (a) of which the RHS of (c) is a particular case. Hence, the direct part of (c) is shown. The reverse part of (c) can be proved by induction on N. For $N = 2$, $R_2^{(N)}(N-1)$ can be shown explicitly using the direct characterization (11). Now, assuming that $R_2^{(N)}(N-1)$ for non-ordered $\boldsymbol{n}$, we would like to show that $R_2^{N+1}(N)$ holds. Let us write

$$\min_{j \geq 0} d_{(n_0,...,n_N)}(j) + d_{(j,n_{N+1})}(0) = \min_{j \geq 0} d_{(\tilde{n}_0,...,\tilde{n}_{i-1},\tilde{n}_{i+1},...,\tilde{n}_{N+1})}(j) + d_{(j,\tilde{n}_i)}(0) \tag{65}$$

$$= \min_{k \geq j \geq 0} d_{(\tilde{n}_0,...,\tilde{n}_{i-1},\tilde{n}_{i+1},...,\tilde{n}_N)}(k) + d_{(k,\tilde{n}_{N+1})}(j) + d_{(j,\tilde{n}_i)}(0) \tag{66}$$

$$= \min_{k \geq j' \geq 0} d_{(\tilde{n}_0,...,\tilde{n}_{i-1},\tilde{n}_{i+1},...,\tilde{n}_N)}(k) + d_{(k,\tilde{n}_i)}(j') + d_{(j',\tilde{n}_{N+1})}(0) \tag{67}$$

$$= \min_{j' \geq 0} d_{(\tilde{n}_0,...,\tilde{n}_N)}(j') + d_{(j',\tilde{n}_{N+1})}(0)$$

$$= d_{(n_0,...,n_{N+1})}(0)$$

where the permutation invariance property is used in (65); $R_3^{(N)}(N-1,k)$ is used in (66) since we assume that $R_2^{(N)}(N-1)$ is trues; $\tilde{n}_i$ and $\tilde{n}_{N+1}$ can be permuted according to $R_2^{(2)}(1)$. Finally,

where we set $M \triangleq N^* - 1$ for simplicity of notation. Obviously, the minimum of the RHS of (71) is achieved with such k^* that

$$k^* - p_M + \left\lfloor \frac{k^* - 1}{M} \right\rfloor \leq 0, \tag{72}$$

$$\text{and } (k^* + 1) - p_M + \left\lfloor \frac{k^*}{M} \right\rfloor > 0. \tag{73}$$

Let us decompose k^* as $k^* = aM + b$ with $b \in [1, M]$. Then, (72) becomes

$$aM + b - p_M + a \leq 0 \tag{74}$$

which also implies that $aN + 1 - p_M + a \leq 0$ from which

$$a = \left\lfloor \frac{p_M - 1}{M + 1} \right\rfloor.$$

The form of a suggests that p_M can be decomposed as

$$p_M = a(M + 1) + \bar{b}. \tag{75}$$

From (74) and (75), we have $b \leq \bar{b}$ and thus $b = \min\{M, \bar{b}\}$. With the form of optimal k and some basic manipulations, we have finally

$$\sum_{i=1}^{p_M} \left(i - p_M + \left\lfloor \frac{i-1}{M+1} \right\rfloor \right) - \sum_{i=1}^{k^*} \left(i - p_M + \left\lfloor \frac{i-1}{M} \right\rfloor \right) = 0$$

which ends the proof.

APPENDIX V

PROOF OF THEOREM 7

It can be proved by showing a stronger result : the asymptotical pdf of $\boldsymbol{\alpha}(\boldsymbol{\Pi}_{\text{PF}}^\dagger \boldsymbol{\Pi}_{\text{PF}})$ in the high SNR regime is identical to that of $\boldsymbol{\alpha}(\boldsymbol{\Pi}^\dagger \boldsymbol{\Pi})$. We show it by induction on N. For $N = 1$, since $\underline{\boldsymbol{H}}_1 = \boldsymbol{H}_1$, the result is direct. Suppose that the theorem holds for N. Let us show that it also holds for $N + 1$. Note that

$$\boldsymbol{\Pi}'_{\text{PF}} = \underline{\boldsymbol{H}}_{N+1} \boldsymbol{P}_N \boldsymbol{\Pi}_{\text{PF}} = \underline{\boldsymbol{H}}_{N+1} \underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^\dagger \boldsymbol{\Pi}_{\text{PF}},$$

from which we have

$$\left(\boldsymbol{\Pi}'_{\text{PF}} \right)^\dagger \boldsymbol{\Pi}'_{\text{PF}} \sim \mathcal{W}_{n_0}(n_{N+1}, (\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^\dagger \boldsymbol{\Pi}_{\text{PF}})^\dagger (\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^\dagger \boldsymbol{\Pi}_{\text{PF}}))$$

$$\sim \mathcal{W}_{\underline{n}_{\min}}(n_{N+1}, \boldsymbol{\lambda}((\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^\dagger \boldsymbol{\Pi}_{\text{PF}})^\dagger (\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^\dagger \boldsymbol{\Pi}_{\text{PF}})))$$

for a given $\mathbf{\Pi}$. Similarly, $\mathbf{\Pi}'^{\dagger}\mathbf{\Pi}' \sim \mathcal{W}_{n_{\min}}(n_{N+1}, \boldsymbol{\lambda}(\mathbf{\Pi}^{\dagger}\mathbf{\Pi}))$. In the high SNR regime, we can show that

$$\boldsymbol{\alpha}((\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}})^{\dagger}(\underline{\boldsymbol{D}}_N \underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}})) = \boldsymbol{\alpha}((\underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}})^{\dagger}(\underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}}))$$
$$= \boldsymbol{\alpha}(\mathbf{\Pi}_{\mathrm{PF}}^{\dagger}\mathbf{\Pi}_{\mathrm{PF}})$$

where the first equality comes from lemma 8 and the second one holds because

$$(\underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}})^{\dagger}(\underline{\boldsymbol{Q}}_N^{\dagger}\mathbf{\Pi}_{\mathrm{PF}}) = \mathbf{\Pi}_{\mathrm{PF}}^{\dagger}\mathbf{\Pi}_{\mathrm{PF}}.$$

Finally, since we suppose that the joint pdf of $\boldsymbol{\alpha}((\mathbf{\Pi}_{\mathrm{PF}}^{\dagger})\mathbf{\Pi}_{\mathrm{PF}})$ is the same as that of $\boldsymbol{\alpha}(\mathbf{\Pi}^{\dagger}\mathbf{\Pi})$, we can draw the same conclusion for $\boldsymbol{\alpha}(((\mathbf{\Pi}_{\mathrm{PF}}')^{\dagger})\mathbf{\Pi}_{\mathrm{PF}}')$ and $\boldsymbol{\alpha}((\mathbf{\Pi}')^{\dagger}\mathbf{\Pi}')$.

REFERENCES

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[2] ——, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[3] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[6] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[7] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[8] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 647–663, Feb. 2007.

[9] ——, "Towards the optimal amplify-and-forward cooperative diversity scheme," Mar. 2006, accepted for publication. [Online]. Available: http://arxiv.org/pdf/cs.IT/0603123

[10] P. Elia and P. V. Kumar, "Approximately universal optimality over several dynamic and non-dynamic cooperative diversity schemes for wireless networks." [Online]. Available: http://fr.arxiv.org/pdf/cs.IT/0512028

[11] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[12] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space-time block codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3885–3902, Dec. 2006.

[13] P. Elia, B. A. Sethuraman, and P. V. Kumar, "Perfect space-time codes with minimum and non-minimum delay for any number of antennas," *IEEE Trans. Inform. Theory*, Dec. 2005, submitted for publication.

[14] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3233–3258, July 2006.

[15] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden code: A 2×2 full-rate space-time code with non-vanishing determinants," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[16] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Diagonal algebraic space time block codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 628–636, March 2002.

[17] S. Alamouti, "Space-time block coding: A simple transmitter diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[18] J. W. C. Jakes, *Microwave mobile comunications*.   New York: Wiley, 1974.

[19] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Math. Statistics*, vol. 35, pp. 475–501, 1964.

[20] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Multivariate Analysis*, vol. 73, pp. 155–165, May 2000.

[21] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

[22] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[23] S. Yang and J.-C. Belfiore, "Diversity-multiplexing tradeoff of double scattering MIMO channels," *IEEE Trans. Inform. Theory*, Mar. 2006, submitted for publication. [Online]. Available: http://arxiv.org/pdf/cs.IT/0603124
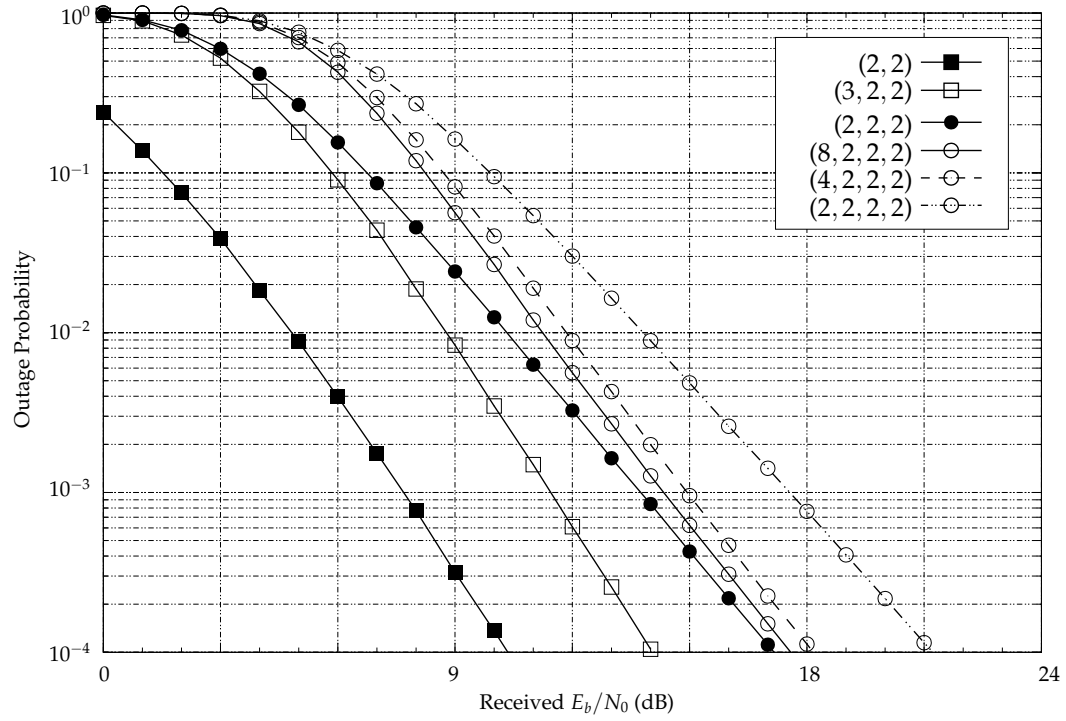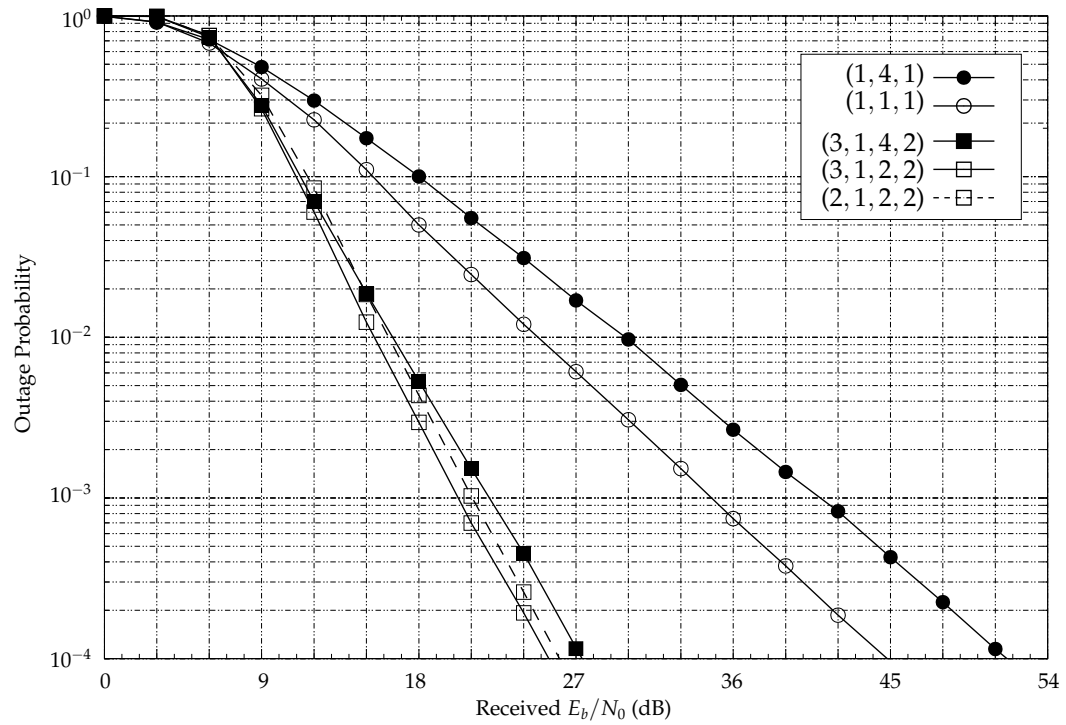
Fig. 7. Horizontal reduction.

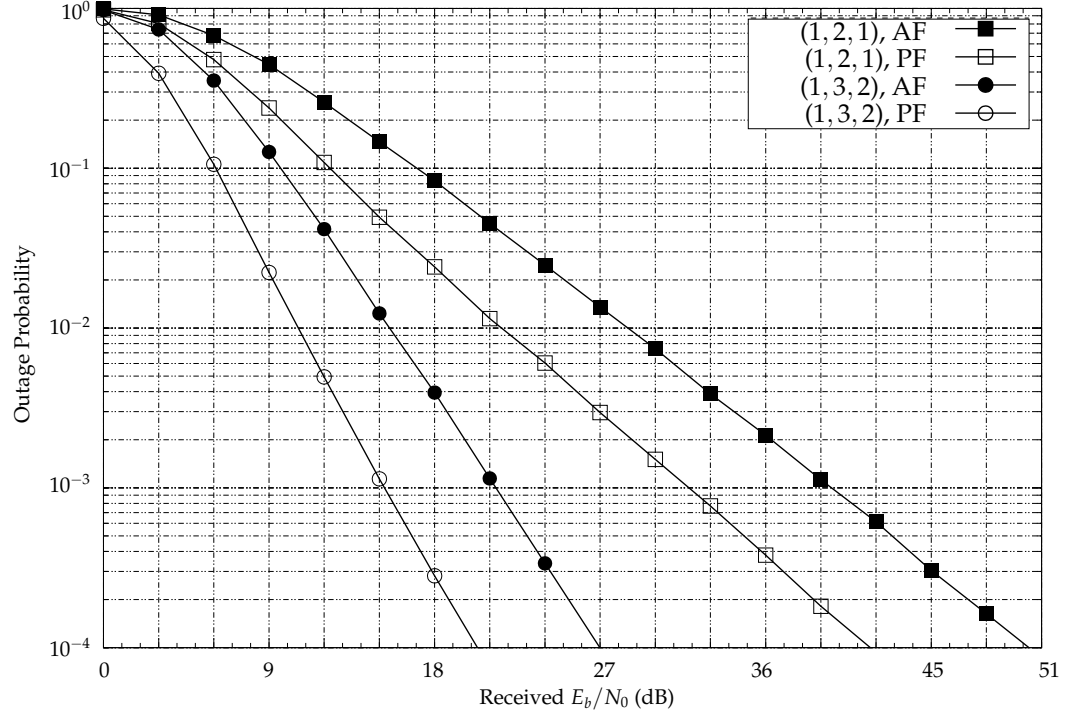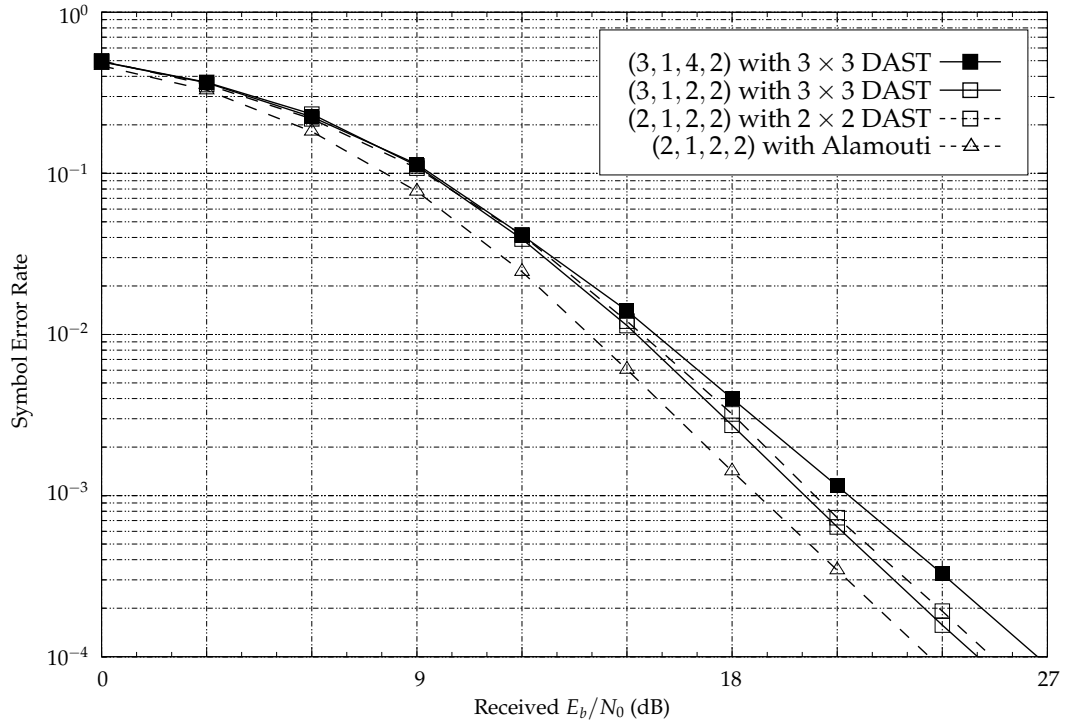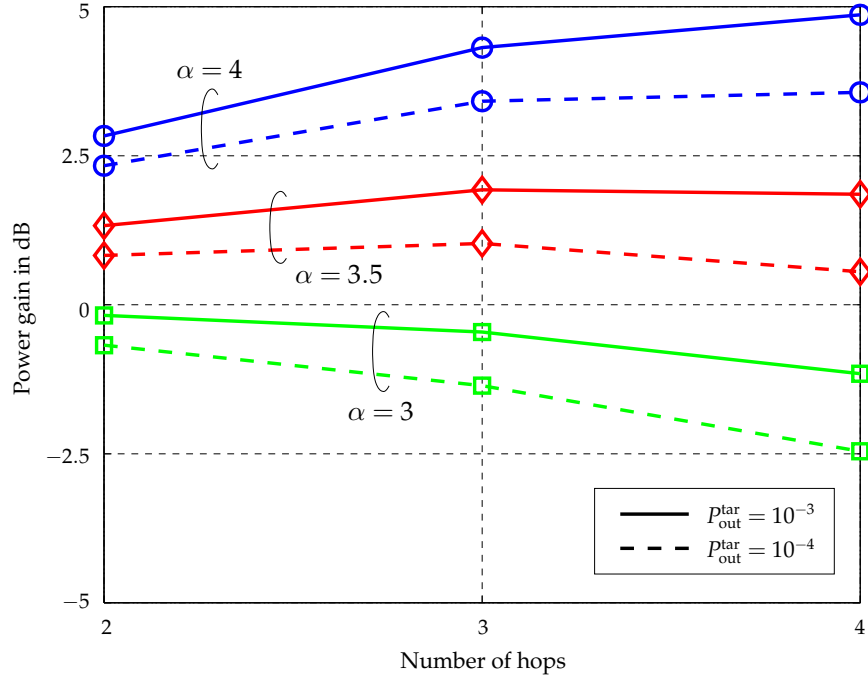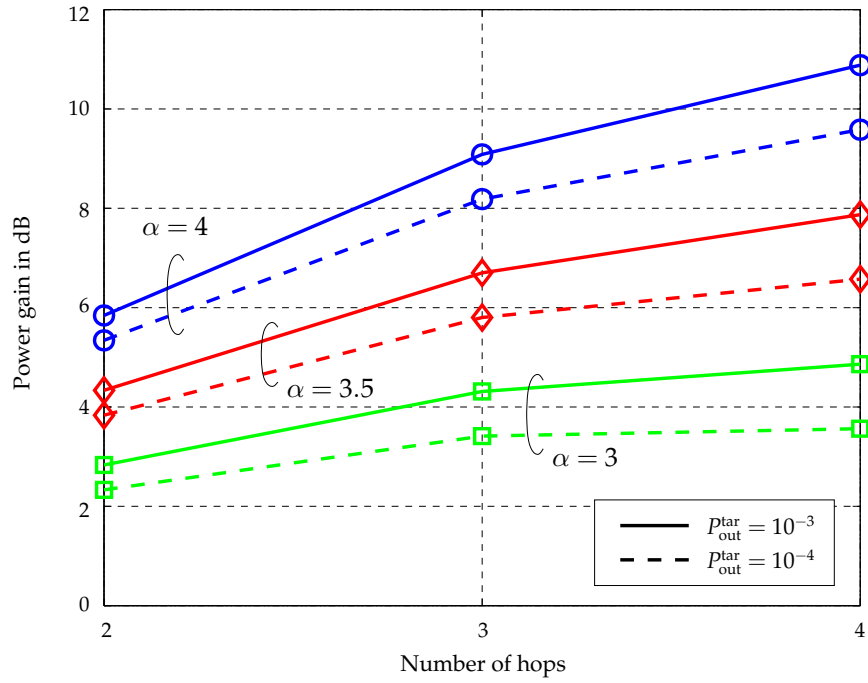Fig. 8. Vertical reduction.

Fig. 9. AF vs. PF.

Fig. 10. Symbol error rate of coded performance.

(a) Total transmission power

(b) Individual transmission power

Fig. 11. Transmission power gain of the AF multihop channel over the direct transmission.